The Neuroscience and Law Center presents

CUTTING-EDGE DEVELOPMENTS IN NEUROSCIENCE AND LAW

February 25, 2020
Skadden Conference Center | Fordham Law School

CLE COURSE MATERIALS
Table of Contents

1. Speaker Biographies (view in document)

2. CLE Materials

Panel 1: Ethics and the Brain: How Environment Affects Behavior and Brain Development
(Supporting Document) Panel 1-3 Intro and Summary to CLE Materials. [View in document]

(Supporting Document 2) Part 1-3 Intro and Summary to CLE Materials. [View in document]


Grady. Christine; Greely, Henry T.; Ramos, Khara M. Neuroethics in the Age of Brain Projects. [View in document]

Panel 2: The Future of Neuroscience and its Moral and Ethical Implications

Lollini, Andrea. Brain Equality: Legal Implications of Neurodiversity in a Comparative Perspective. [View in document]

Chapter 3 - Neuroethics Roadmap. NEUROETHICAL ISSUES AND NEUROTECHNOLOGIES. [View in document]

Panel 3: The Ethics of Trauma, Aging, and Brain Irregularities

Chapter 6 - Neuroethics Roadmap. INTEGRATING NEUROETHICS AND NEUROSCIENCE. [View in document]

Rosenthal, Hilary. SCANNING FOR JUSTICE: USING NEUROSCIENCE TO CREATE A MORE INCLUSIVE LEGAL SYSTEM. [View in document]

Ienca, Marcello; Andorno, Roberto. Towards new human rights in the age of neuroscience and neurotechnology. [View in document]

Panel 4: The Law and Ethics of Neuroscientific Development
(Supporting Document) Panel 4 Intro and Summary to CLE Materials. [View in document]


Panel 1 Additional Materials

Raine, Adrian; Glenn, Andrea L. Perspectives: Neuroscience and the Law - Science and Society. Neurocriminology: implications for the punishment, prediction and prevention of criminal behaviour. [View in document]

Raine, Adrian; Leung, Chi-Ching; Singh, Melvinder; Kaur, Jasmin. Journal of Experimental Criminology. Omega-3 supplementation in young offenders: a randomized, stratified, double-blind, placebo controlled, parallel-group trial. [View in document]

Choy, Olivia; Raine, Adrian; Hamilton, Roy H.. The Journal of Neuroscience. Stimulation of the Prefrontal Cortex Reduces Intentions to Commit Aggression: A Randomized, Double-Blind, Placebo-Controlled, Stratified, Parallel-Group Trial. [View in document]

Panel 2 Additional Materials


Panel 4 Additional Materials

Choudhary, Arabinda Kumar; Sarvaes, Sabah; Slovis, Thomas L.; Palusci, Vincent J.; Hedlund, Gary L.; Narang, Sandeep K.; Moreno, Joelle Anne; Dias, Mark S.; Christian, Cindy W.; Nelson Jr., Marvin, D.; Silvera, V. Michelle; Palasis, Susan; Raisaki, Maria; Rossi, Andrea; Offiah, Amaka. Consensus statement on abusive head trauma in infants and young children.
Introduction

These CLE materials were gathered from a 2019 report titled the “Neuroethics Roadmap,” which was created by the BRAIN initiative, a group established under the Obama administration. This memo briefly describes some of the objectives and goals of the BRAIN Initiative, then summarizes three of the report’s chapters (Chapters 1, 3, and 6 assigned for the first three symposium panels respectively). The summary emphasizes in particular discussions concerning ethics in neuroscientific research and the elimination of bias.

Neuroethics and the Brain Initiative

The White House BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies), is a collaborative, public-private research initiative announced by the Obama administration on April 2, 2013.¹ BRAIN was established with the goal of supporting the development and application of innovative technologies that can create a dynamic understanding of brain function. By accelerating the development and application of innovative technologies, researchers are more able to produce a revolutionary new dynamic picture of the brain that shows how individual cells and complex neural circuits interact in both time and space. This groundbreaking research allows researchers to understand new ways to treat, cure, and even prevent brain disorders, and hopes to provide greater knowledge on how the brain enables the human body to record, process, utilize, store, and retrieve vast quantities of information, all at the speed of thought. The Initiative unites federal agencies, private foundations, as well as teams of neuroscientists in order to advance knowledge concerning the brain’s billions of neurons and gain greater insights into perception, actions and, ultimately, consciousness.

Chapter 1: NEUROETHICS PAST, PRESENT, AND FUTURE

The first chapter of the Neuroethics Roadmap examines findings and analyses regarding neuroethics issues in current and potential research as part of the BRAIN Initiative. The chapter declares that brain disease impacts every corner of the globe, but also emphasizes how those impacts “are worsened through health disparities that limit availability of effective treatments to all individuals who could benefit from them.” This highlights the potential for bias in neuroethics.

This chapter also considers how insights in neuroscience “are likely to change how we perceive ourselves as individuals and as members of society.” As a result, the Roadmap argues, we must discern how this newfound knowledge will be utilized not only in the classroom, but also

in instances such as “the validity of neuroscience measurements for judging intent or accountability in our legal system, the use of neuroscience insights to mount more persuasive advertising or public service campaigns, the issue of privacy of one’s own thoughts and mental processes in an age of increasingly sophisticated neural ‘decoding’ abilities, and many other questions.” It emphasizes how there is both tremendous opportunity for important research, but with that also arises a “vital role of protecting research participants and guiding against potential malign intent by rogue actors.”

Neuroscience has the potential to “reveal core mechanisms that underlie human thoughts, emotions, perceptions, actions, identity, and memories. Ethical questions and challenges are intertwined with this research, the results of which may change what many people view as consciousness, agency, and human nature.” This amount of potentially sensitive information requires an examination of the ethical policies surrounding neuroscience. These policies, principles, and values are referred to as “neuroethics.” “Neuroethics is concerned with ethical, legal, and social issues arising from the conduct, application, and implications of neuroscience and neurotechnologies in a variety of contexts. The International Neuroethics Society defines neuroethics as ‘... a field that studies the implications of neuroscience for human self-understanding, ethics, and policy.’” Issues often emphasized in these kinds of bioethical studies include the ethics of research on animals, data privacy, risk mitigation, and health-care access. Other topics of consideration include privacy, fairness, liberty, personal identity, informed consent, and moral responsibility.

The NIH BRAIN Neuroethics Working Group suggested eight neuroethics guiding principles for Institutional Review Boards and others involved in the conduct of BRAIN Initiative-funded research. The Neuroethics Guiding Principles are meant to instruct neuroscientists, particularly BRAIN Initiative-supported researchers, to help them consider the ethical, legal, and societal implications of their work. These principles include: (1) making assessing safety paramount, (2) anticipating special issues related to capacity, autonomy, and agency, (3) protecting the privacy and confidentiality of neural data, (4) attending to possible malign uses of neuroscience tools and neurotechnologies, (5) using caution when moving neuroscience tools and neurotechnologies into medical or non-medical uses, (6) identifying and addressing specific concerns of the public about the brain, (7) encouraging public education and dialogue, and finally (8) behaving justly and share the benefits of neuroscience research and resulting technologies.

Chapter 1 concludes by arguing that frameworks such as these are essential in making decisions concerning how not only scientists, but we as a society, obtain and utilize this kind of information. These kinds of neurotechnologies have the potential to greatly affect agency, identity, capacity, and public trust, and risks associated with augmentation, hype, bias, and possible misuse of technologies and data. However, the utilization and integration of this type of knowledge could also greatly benefit scholarship concerning philosophy, psychology, law, theology, sociology, and other areas and “have a powerful, positive impact on research and the knowledge it generates.”

Chapter 3: NEUROETHICAL ISSUES AND NEUROTECHNOLOGIES

Chapter 3 of the Roadmap further explores neuroethical implications as new neurotechnologies are discovered and refined. This chapter asserts that if proper consideration is
given to the ethical considerations of advances in neuroscience, “neuroethics may be used not to impede, but to advance the ability of BRAIN research to have the greatest societal impact.” For instance, some of the questions that neuroscientists must consider are: what are the unintended consequences on social stigma and self-stigma? And what are the possible social or cultural biases in research design or interpretation of scientific results?

The chapter uses the eight guiding principles, listed above, to “offer a framework for prioritizing values and for ethical guidance for the conduct of BRAIN research including new technology development.” However, they also identified several “priority areas” as a “framework for identifying companion neuroethical issues and neuroethics research opportunities.” One of these priority areas includes “Discovering Diversity,” which states that the study of neuroscience “should not focus on a single species, or one societal group or sex or race or age, ensuring that bias is mitigated, and that the benefits of neuroscience research can apply to individuals from numerous populations.”

Another priority area is “Maps at Multiple Scales,” which the Initiative suggests means that as neuroscience becomes even more sophisticated, we must consider how the “wider availability of such techniques for human use may prompt non-medical, commercial, consumer, or judicial use of such technologies and care will be needed to determine which contexts technology or innovation can be justly deployed.” This priority area also highlights, again, how in past scientific research the use of male-only samples, for example, have resulted in incorrect conclusions, underscoring the need for a diverse range of subjects that are “fully representative cross section of our society.” The idea of diversity is again raised in the “Brain in Action” priority area, where the Roadmap warns that there is a strong tendency to differentiate between what is “normal” and what is not, and expresses that careful considerations will be needed “related to risks of reinforcing bias by dividing participants along socially constructed identities as well as with privacy.”

Another priority area, “Identifying Fundamental Principles,” discusses the potential ethical issues of data sharing. Data sharing includes things such as analyses, algorithms, and shared access to infrastructure. This is especially important because it implicates privacy issues with regard to identity, and human brain data has the potential to give tremendous insight into an individual’s thoughts and other [private] aspects of an individual.” This priority also emphasizes how studies need to be designed to “investigate the impact of many variables on brain function, including but not limited to, sex, race, and cultural experiences,” paying special attention to who will benefit from these scientific advances and how we can promote “equitability across these and other important domains.”

**Chapter 6: INTEGRATING NEUROETHICS AND NEUROSCIENCE**

Chapter six of the Neuroethics Roadmap considers what steps should be taken in order to integrate neuroethics into the study and practice of current and future neuroscience research. The Roadmap once again emphasizes the importance of neuroethics because it “helps guide neuroscience advances and discoveries toward positive social outcomes – in medical or non-medical settings,” the impacts of which “have significant and broad-reaching ethical implications.” This chapter also emphasizes the importance of integrating neuroethics at every step of a
neuroscience research project, from the very first hypothesis until the project’s conclusion. The Roadmap argues that not only will this integration be beneficial to neuroscientists, but it will help accelerate neuroscience because it will allow scientists to identify and mitigate potential roadblocks which might otherwise arise. For example, by utilizing neuroethical guidelines scientists might consider how the questions they choose to study in the lab might amplify existing biases, such as the all-male pool of participants example used above, which would lead to incorrect conclusions.

Not only can neuroscientists benefit from the use of neuroethics, but ethicists are also benefitted when they are apprised on the latest neuroscience research because they are given “a better sense of how to navigate the ethical challenges.” This chapter also discusses when and how neuroethics might be integrated into neuroscientific research, as well as the importance of public involvement and engagement at all levels of neuroethical studies. This includes the need for scientists to be transparent with their research in order to receive support for their research from the public at large, which hinges on scientists being adept with communicating with the public – especially considering the nature of the media today (and their tendency to embellish or sensationalize information.) The Roadmap argues that neuroscientists should learn from past experiences, such as human-genome editing, where scientists later opined that “[w]e need to engage the public in a more open and honest bidirectional dialogue about science and technology and their products, including not only their benefits but also their limits, perils, and pitfalls. We need to respect the public’s perspective and concerns even when we do not fully share them, and we need to develop a partnership that can respond to them.” By considering all of the potential ethical implications mentioned above and throughout the Roadmap, neuroscientists can better work with neuroethicists together to integrate and implement their shared goals.
Introduction

These CLE materials derive from an article by Bernice Donald and Erica Bakies that was part of a 2016 Neuroscience and Law Center Symposium in the *Fordham Law Review* titled “Criminal Behavior and the Brain: When Law and Neuroscience Collide.”¹ This memo briefly summarizes the article, emphasizing in particular discussions concerning the ethics of neuroscience in the courtroom and the elimination of bias. This reading pertains to the fourth symposium panel.

**Understanding the Role of Neuroscience in Criminal Sentencing**

The article, *A Glimpse Inside the Brain’s Black Box: Understanding the Role of Neuroscience in Criminal Sentencing*, analyzes how advances made within the study of neuroscience have resulted in immense developments in the legal community and criminal justice system—specifically in the context of sentencing. The article begins by discussing what neuroscience and the smaller associated field of study, neuropsychology, are and what they can tell us about an individual. It then recounts a brief history of sentencing in the United States. Additionally, the article expounds on how the legal system currently utilizes neuroscience in the courts, noting specifically the ways in which neuroscience can be presented during the sentencing phase of trial. Finally, the article discusses the use of neuroscience as a mitigating factor during sentencing and how judges can use neuroscience to combat their implicit biases.

The authors approach the issue of ethics by examining how neuroscience can eliminate implicit bias in the criminal justice system. They establish that there are historically large differences in the length of sentences on the basis of race, gender, education, income, and citizenship. They then argue that the utilization of neuroscience has the potential to mitigate those kinds of biases, and help make strides towards leveling the playing field within the criminal justice system. Therefore, it is important for all participants in the criminal justice system, from judges and defense attorneys to prosecutors and legislators, to be informed of the developments in neuroscience to ensure that we are guarding against bias and advocating for fairness and justice.

---

Neuroethics Roadmap

The ACD Working Group on BRAIN 2.0 Neuroethics Subgroup (BNS) was formed to develop a Neuroethics Roadmap for the NIH BRAIN Initiative; review priority areas identified in the BRAIN 2025 Strategic Plan, incorporating updates from the broader BRAIN 2.0 Working Group; and characterize the neuroethical implications that may arise as BRAIN Initiative investments produce new tool/neurotechnologies, and/or those tools/neurotechnologies are applied to advancing the goals of the NIH BRAIN Initiative.

The BNS has conducted a portfolio review and held a public workshop on neuroethical issues posed by research through the BRAIN Initiative. The BNS would now appreciate comments from the public on draft findings and analysis detailed in a Neuroethics Roadmap. The Neuroethics Subgroup has also provided analysis and findings to the Working Group on BRAIN 2.0 for inclusion in the Working Group's initial thoughts. The public comment period on the Neuroethics Roadmap has closed. The findings and analysis from the BNS will be presented to the Advisory Committee to the Director, NIH, for consideration at a public meeting on Jun 13-14, 2019.

THE BRAIN INITIATIVE AND NEUROETHICS:
ENABLING AND ENHANCING NEUROSCIENCE ADVANCES FOR SOCIETY

CHAPTER 1. NEUROETHICS PAST, PRESENT, AND FUTURE

The Brain Research through Advancing Innovative Neurotechnologies® (BRAIN) Initiative aims to revolutionize our understanding of the human brain with a priority of developing and using new tools and neurotechnologies for “acquiring fundamental insight about how the nervous system functions through understanding circuit function from single cells to complex behaviors.” The BRAIN Initiative conveys an ethical imperative to improve human health through building an understanding of the brain and its functions. Brain disorders exert a considerable toll on human health, accounting for a substantial proportion of the world's health problems. Together, brain injury combined with neurologic, psychiatric, and substance-use disorders are a leading cause of the global burden of disease, and their incidence is expected to rise. Yet, effective methods to prevent and treat brain disorders remain limited, despite substantial investments in neuroscience research.

About this Document

In this Neuroethics Roadmap, the NIH ACD BRAIN Initiative Neuroethics Subgroup (BNS) presents its findings and analyses regarding neuroethics issues in current and potential research as part of the BRAIN Initiative. The BNS also offers some specific suggestions regarding NIH activities in the BRAIN Initiative. The BNS proposes that the Advisory Committee to the Director, NIH (ACD) recommend to the NIH Director that the NIH, specifically the NIH BRAIN Initiative, consider the findings, analysis, and suggestions in this report for integration, rather than as a parallel effort, into the BRAIN Initiative research program. The BNS recognizes that the some of their findings and suggestions go beyond the mission of the NIH or would require NIH to work with other Federal agencies and non-Federal entities and stakeholders. In those cases, the BNS proposes that the ACD recommend to the NIH Director that NIH look for opportunities to engage with broader stakeholder communities to address issues and achieve outcomes consistent with the spirit of the BNS Neuroethics Roadmap.
Brain diseases affect people across the lifespan and in every corner of the globe, exerting considerable impacts on society, which affect public health and economic stability. These impacts are worsened through health disparities that limit availability of effective treatments to all individuals who could benefit from them. BRAIN Initiative-generated neurotechnologies and insights are transforming our capacity to understand complex spatial and temporal circuits and systems, thereby offering new hope to prevent, diagnose, or treat devastating brain diseases that rely on proper function of these circuits.

The first 5 years of this unparalleled effort has yielded significant discoveries in all seven priority areas designated by the BRAIN Initiative’s flagship strategic plan, BRAIN 2025: A Scientific Vision (see text box above) (https://www.braininitiative.nih.gov/strategic-planning/brain-2025-report). These advances span three-dimensional maps of cell types and activity-dependent gene expression, high-speed three-dimensional imaging of neural activity, novel methods of neuromodulation, a range of sensors and probes that continue to advance ongoing discovery, among others. Powerful new modes of computational analyses and data-science methods such as machine learning offer powerful and efficient investigational tools. Although neuroscience research entails ethical issues common to other areas of biomedical science, it conveys other, unique considerations. From its beginning, the BRAIN Initiative appreciated the importance of neuroethics, as articulated in the BRAIN 2025 report:

“… mysteries unlocked through the BRAIN Initiative, and through neuroscience in general, are likely to change how we perceive ourselves as individuals and as members of society. Many of these discoveries will raise more questions than they answer. We may need to consider, as a society, how discoveries in the area of brain plasticity and cognitive development are used to maximize learning in the classroom, the validity of neuroscience measurements for judging intent or accountability in our legal system, the use of neuroscience insights to mount more persuasive advertising or public service campaigns, the issue of privacy of one’s own thoughts and mental processes in an age of increasingly sophisticated neural ‘decoding’ abilities, and many other questions. Questions of this complexity will require insight and analysis from multiple perspectives and should not be answered by neuroscientists alone.”

**BRAIN 2025: 7 Priority Areas**

- Priority Area 1. Identify and provide experimental access to different brain cell types to determine their roles in health and disease
- Priority Area 2. Generate circuit diagrams ranging from synapses to the whole brain
- Priority Area 3. Develop and apply improved methods to monitor neural activity
- Priority Area 4. Link brain activity to behavior through precise interventional tools that change neural-circuit dynamics
- Priority Area 5. Understand the biological basis of mental processes via new theoretical and data-analysis tools
- Priority Area 6. Develop innovative technologies to understand the human brain and treat its disorders
- Priority Area 7. Integrate new technological and conceptual approaches produced in goals #1-6 toward understanding cognition, emotion, perception, and action in health and disease

**The role of neuroethics in the BRAIN Initiative**

BRAIN Initiative research also has exciting but unknown potential to challenge the typical ways we think about life, death, each other, and ourselves. The results of neuroscience research affect what we know about how the brain produces complex thoughts and behaviors. This knowledge could reveal core mechanisms that underlie human thoughts, emotions, perceptions, actions, identity, and memories. Ethical questions and challenges are intertwined with this research, the results of which may change what many people view as consciousness, agency, and human nature. It is clear that as our understanding of the brain increases – along with the arrival of neurotechnologies to intervene with its many functions – neuroethical questions will likely emerge at every turn. These questions deserve sensitive and systematic responses, as well as development of concrete implementable goals to ensure that neuroscience research and neuroethics are tightly integrated.

The authors of the BRAIN 2025 report also noted, “Although brain research entails ethical issues that are common to other areas of biomedical science, it entails special ethical considerations as well. Because the brain gives rise to consciousness, our innermost thoughts and our most basic human needs, mechanistic studies of the brain have already resulted in new social and ethical questions.” Recognizing the vitality of ethics to the BRAIN Initiative, the National Institutes of Health (NIH) established a Neuroethics Working Group (https://www.braininitiative.nih.gov/about/neuroethics-working-group) to anticipate and recommend overall
approaches for how the BRAIN Initiative might navigate ethical issues. The Neuroethics Working Group has held several public meetings and published guiding principles and other papers in the neuroethics literature. Since 2017, the NIH BRAIN Initiative has also issued specific neuroethics funding announcements and has funded neuroethics research projects.

In the recent, interim review of BRAIN 2025 (“BRAIN 2.0”), the NIH Director re-emphasized the integral value of neuroethics in the BRAIN Initiative and established a neuroethics subcommittee of the Advisory Committee to the Director (ACD) conducting this mid-course review. He charged this group, the NIH ACD BRAIN Initiative Neuroethics Subgroup (BNS) with developing a Neuroethics Roadmap (this document) and called for integration of neuroethics in future BRAIN-Initiative efforts by incorporating neuroethics principles and research opportunities into the BRAIN 2.0 report, work that is underway.

What is neuroethics and why is neuroethics important?

Bioethics involves deliberation, analysis, and research to inform well-considered ethical decisions guided by principles and values as well as notions of fairness, respect, and benefit. Bioethical principles can help guide and shape practices, institutions, implications, and social impacts of science, clinical practice, and health care. While certain bioethical issues frequently recur in biomedicine and clinical practice, others emerge anew as science evolves. Among current topics of considerable interest within the realm of bioethics are synthetic biology, gene-editing, and transplantation; many issues related to clinical research and care; and issues related to research with animals and humans.

More specifically, neuroethics is concerned with ethical, legal, and social issues arising from the conduct, application, and implications of neuroscience and neurotechnologies in a variety of contexts. These range from medical practice to research to commercial interests to society at large. The International Neuroethics Society defines neuroethics as "... a field that studies the implications of neuroscience for human self-understanding, ethics, and policy." There are many issues important to neuroethics that are familiar and important bioethical issues. These include the responsible conduct of research, the ethics of research with humans and non-human animals, data privacy, risk mitigation, health-care access, and others. Yet because brain function is intimately connected to our understanding of identity, human responsibility, privacy, authority, agency, personhood, and normality, neuroethics is distinct from the broader topic of bioethics. Neuroethics also has an important role in sorting the opportunities and ramifications of various neurotechnologies – defined as any technology that informs our understanding of the brain and its functions, including higher-order activities like consciousness and thought. Neurotechnologies are currently being developed as both research tools (to visualize or otherwise measure brain function) and as therapies, to repair brain dysfunction. (see Chapter 3. Neuroethics and Neurotechnologies)

At two extremes, perceptions of neuroethics may appear esoteric or even punitive. Importantly, neuroethics is not a set of rules or compliance mechanisms, and its role should not be seen as limited to implementing oversight of the responsible conduct of research. Rather, fully integrating neuroethics with neuroscience offers tremendous opportunity for new research insights, inviting new fields including the humanities into scientific discourse, bringing science to people in a way they care about it – in addition to its vital role of protecting research participants and guarding against potential malign intent by rogue actors.

Neuroethics encompasses reflection, analysis, and research – to inform, enable, and strengthen neuroscience research. The intended reach of neuroethics goes beyond ethical conduct of neuroscience research, to the clinical and societal applications of this work. Neuroethical consideration frames responsible acquisition and use of knowledge about the brain and the nervous system. It also facilitates planning for – and in some cases, adjusting for the implications of – how such knowledge is applied to human health, illness, and behavior. Given the unprecedented precision of new neurotechnologies and the brain’s centrality to human identity, familiar bioethics topics take on new dimensions and complexities. These include privacy, fairness, liberty, personal identity, informed consent, and moral responsibility. Neuroethics might tackle questions such as: i) which brain circuits or function influence our ability to act rationally to be capable of voluntary, intentional actions; ii) what is authenticity, and is it jeopardized when our executive function is damaged or when an implanted central nervous system device alters our interests, evaluations, or responses?; iii) when are people not responsible for their actions and behavior, and/or do certain neurological characteristics or neural devices reduce culpability for their actions?; iv) what does privacy mean in the setting of neurotechnologies, and how does one protect against possible threats to people’s innermost thoughts?; and v) how should one apply considerations of justice related to neural development, plasticity, and access to possible technological improvements. Many more questions are easily conceivable. These topics are considered throughout the chapters of this Neuroethics Roadmap, both in the context of current research and as they relate to future discoveries and technologies.

What existing ethical guidance applies to neuroethics and is there a need for more?
Multiple sources of guidance and guidelines currently inform neuroscience research; these are described briefly below and do not constitute an exhaustive analysis but rather highlight recent useful approaches to thinking about and dealing with neuroethical issues. Using these and other guidelines and methods, however, there remains a need for deliberation and dialogue, anticipating possible impacts on individuals, populations, and society (see Chapter 6. Integrating Neuroethics).

**Belmont Report**

Often considered a seminal source of ethics guidance, the 1978 *Belmont Report* (https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html), issued by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, is somewhat unique among ethical guidance in that it concisely analyzes three principles that underlie the ethics of research with humans (see Table 1). The report applies these principles – respect for persons, beneficence, and justice – to specific research-related activities. The Belmont principles are known for their simplicity, clarity, reach, and endurance – but they require context and balancing when applied to individual applications. They form the basis for U.S. federal regulations (both the Common Rule (https://www.hhs.gov/ohrp/regulations-and-policy/regulations/finalized-revisions-common-rule/index.html) and Food and Drug Administration (FDA) regulations (https://www.fda.gov/scienceresearch/specialtopics/runningclinicaltrials/ucm155713.htm)) and other ethical-guidance documents that govern protection of human research participants. Importantly, however, the three Belmont principles are not unique to clinical research – they are familiar and applicable principles in other domains, such as health care. Over time, potential limitations of the Belmont principles have been raised (https://www.ncbi.nlm.nih.gov/pubmed/?term=28661753) in light of evolving research practices and the importance of other considerations such as transparency and the impact of research on groups. This debate has led to rethinking (https://www.ncbi.nlm.nih.gov/pubmed/?term=16036651) possible additional principles for the ethical conduct of research. More recently, some have suggested the need for a set of Belmont principles specific to research with neurotechnologies (https://linkinghub.elsevier.com/retrieve/pii/S0092-8674(16)31449-0) and to neuroscience.

**Table 1. BELMONT PRINCIPLES**

*National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research 1978*

<table>
<thead>
<tr>
<th>Principle</th>
<th>Explanation</th>
<th>Application to clinical research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect for persons</td>
<td>To respect, and not interfere with, the self-determined choices and actions of autonomous individuals; and to provide additional protections for those with diminished autonomy</td>
<td>Informed consent for enrollment and ongoing participation</td>
</tr>
<tr>
<td>Beneficence</td>
<td>To never deliberately harm another, to maximize benefits and minimize risks, and to promote the welfare of others</td>
<td>Analysis of risks and benefits and determination that benefits justify the risks</td>
</tr>
<tr>
<td>Justice</td>
<td>To be fair in the distribution of social goods such as the benefits and burdens of research</td>
<td>Fair procedures and outcomes in the selection of subjects</td>
</tr>
</tbody>
</table>

**Presidential Commission for the Study of Bioethical Issues**

At the outset of the BRAIN Initiative in 2013, the Presidential Commission for the Study of Bioethical Issues was tasked to identify “…proactively a set of core ethical standards – both to guide neuroscience research and to address some of the ethical dilemmas that may be raised by the application of neuroscience research findings.” The Commission sought the advice of many experts at multiple public hearings and published two volumes entitled *Gray Matters* (https://bioethicsarchive.georgetown.edu/pcsbi/node/4704.html). The first recommends the integration of ethics early and explicitly throughout the processes of neuroscience research. Such integration could take several forms, such as education at all levels; institutional infrastructure; research on the ethical, legal, and social implications of BRAIN Initiative research; consultation on research ethics; stakeholder engagement; and inclusion of an ethics perspective within the research team. The second volume recognized that while some ethical issues in neuroscience are not unique to neuroscience, many become even more pronounced. *Gray Matters, Volume 2*
focused on three controversial and timely topics that illustrate ethical tensions and societal implications: cognitive enhancement, consent capacity, and neuroscience and the legal system, and issued 14 specific recommendations across these three areas (see text box, below).

**Gray Matters Vol. 2 Neuroethics Recommendations for Neuroscience**

1. Prioritize Existing Strategies to Maintain and Improve Neural Health  
2. Prioritize Treatment of Neurological Disorders  
3. Study Novel Neural Modifiers to Augment or Enhance Neural Function  
4. Ensure Equitable Access to Novel Neural Modifiers to Augment or Enhance Neural Function  
5. Create Guidance About the Use of Neural Modifiers  
6. Responsibly Include Participants with Impaired Consent Capacity in Neuroscience Research  
7. Support Research on Consent Capacity and Ethical Protections  
8. Engage Stakeholders to Address Stigma Associated with Impaired Consent Capacity  
9. Establish Clear Requirements for Identifying Legally Authorized Representatives for Research Participation  
10. Expand and Promote Educational Tools to Aid Understanding and Use of Neuroscience within the Legal System  
11. Fund Research on the Intersection of Neuroscience and the Legal System  
12. Avoid Hype, Overstatement, and Unfounded Conclusions  
13. Participate in Legal Decision-Making Processes and Policy Development  
14. Establish and Fund Multidisciplinary Efforts to Support Neuroscience and Ethics Research and Education

In addition to the Gray Matters reports, the Commission had previously published a set of principles useful for assessing emerging technologies, appearing in *New Directions: The Ethics of Synthetic Biology and Emerging Technologies* (https://bioethicsarchive.georgetown.edu/pcsbi/synthetic-biology-report.html). These principles are intended to illuminate and guide public policy choices to ensure that new technologies, including synthetic biology, are developed in an ethically responsible manner (see Table 2 below). They complement other sources of ethical guidance and are relevant to development and application of new neurotechnologies and BRAIN Initiative research.

**Table 2. Principles for Assessing Emerging Technologies**  
**Presidential Commission for the Study of Bioethical Issues 2010**

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Beneficence</td>
<td>Responsibility to maximize public benefits while minimizing public harms</td>
</tr>
<tr>
<td>Responsible Stewardship</td>
<td>Prudent vigilance- practical, sensible, cautious ways for assessing likely benefits, safety, and security risks both before and after projects are undertaken. Limiting scientific projects and exploration when necessary out of collective concern for current and future people and the environment</td>
</tr>
<tr>
<td>Intellectual Freedom and Responsibility</td>
<td>Intellectual freedom coupled with the responsibility of individuals and institutions to use their creative potential in morally responsible ways.</td>
</tr>
<tr>
<td>Democratic Deliberation</td>
<td>Collaborative decision making that embraces respectful debate of opposing views and active participation of citizens and the public</td>
</tr>
<tr>
<td>Justice and Fairness</td>
<td>Concern about fair distribution of the benefits and burdens across society</td>
</tr>
</tbody>
</table>

*Nuffield Council on Bioethics*

**Table 3. Ethical Framework for Novel Neurotechnologies**  
**UK Nuffield Council 2013**
Foundational Principles

Beneficence - developing interventions and technologies to address suffering from brain disorders, yet

Caution - because of uncertainties and possible effects on our brains

Key Interests

Individual interests in safety, protection risks, impacts on privacy, and promotion of autonomy

Public interests in equity of access and promoting trust in neurotechnologies

Virtues - necessary in promoting and protecting the identified interests

In Neurotechnologies: Intervening in the Brain (http://nuffieldbioethics.org/wp-content/uploads/2013/06/Novel_neurotechnologies_report_PDF_web_0.pdf), the Nuffield Council on Bioethics – an independent body in the United Kingdom that examines and reports on ethical issues in biology and medicine – proposed two foundational ethical principles (beneficence and caution), as well as key interests and virtues that together form an ethical framework. Beneficence is required for developing and applying therapeutic neurotechnologies because of the "... suffering caused by brain disorders and an absence of other suitable treatments," but caution is also needed based upon uncertainty about the benefits and risks of these technologies, their novelty, and possible unexpected effects given our still-limited knowledge of brain function. In articulating implications of the two principles, the Council identified five interests warranting particular attention for individuals (for effects of treatment decisions on people's lives) and to the public more generally. The five key interests are: i) protection of safety, taking into account risks and expected benefits; ii) promotion of autonomy, in the sense of supporting an individual's capacity to make his or her own decisions; iii) protection of individual privacy, bearing in mind that some devices may collect sensitive personal data; iv) promotion of equity both in terms of access to innovative products and for addressing social stigma and discrimination; and v) promoting public understanding of, and trust in, novel neurotechnologies. Finally, this Council proposed that in seeking to protect and promote these identified interests, three virtues are especially relevant and should guide the activities of all involved parties across a wide range of settings and technologies. These are: i) inventiveness (such as technological innovation) and identifying ways to enhance access; humility in acknowledging the limits of our knowledge and capabilities; and responsibility, through robust research and clinical practices and avoiding exaggeration, hype, or premature claims.

Morningside Group

Concerned that existing guidance was insufficient or not specific enough to address the complex issues presented by neurotechnologies focused upon brain-computer interfaces and artificial intelligence, a multidisciplinary team convened meetings in 2016 and 2017, which resulted in the "Morningside Group" guidance. This group of neuroscientists, neurotechnologists, clinicians, ethicists, and machine-intelligence engineers identified four major distinct ethical issues related to neurotechnologies and artificial intelligence (https://www.researchgate.net/publication/326382606_Will_artificial_intelligence_be_a_blessing_or_concern_in_assistive_robots_for_play): i) privacy and consent, ii) agency and identity, iii) augmentation, and iv) bias. They offered several recommendations to address these concerns globally, including adding "neuro-rights" to international treaties, regulating the use of neurotechnology for augmentation and military use, and regulating the use and sale of neural data.

Global Neuroethics Working Group of the International Brain Initiative

Held annually in South Korea since 2017, the Global Neuroethics Summit (https://globalneuroethicssummit.com/), a workshop hosted by the Neuroethics Workgroup of the International Brain Initiative (https://www.internationalbraininitiative.org/), links global neuroethics efforts around the globe. Leveraging momentum from an international consortium of seven large-scale nation-level brain-initiative efforts, the Summit recognized the critical influence of cultural values and perspectives related to both neuroethics and neuroscience – in particular, highlighting the need for culturally informed and culturally aware neuroethical inquiry. Summit delegates have developed a set of cross-cultural neuroethics questions meant to encourage neuroscientists across various brain projects to consider neuroethical questions (NeQN (https://www.cell.com/neuron/fulltext/S0896-6273(18)30823-7), see Table 4). The questions are further discussed and applied throughout this Neuroethics Roadmap.
The NeQNs are intended to be adaptable and informed by country-relevant cultural values and frameworks, with the goal of acknowledging possible diverse understandings and values related to specific concepts and interests. For example, the need to protect brain-research participant privacy is universally important but varies in scope. Summit delegates concluded that developing a culturally informed global framework for neuroethics requires attention to inclusivity, education, and communication. These questions were used by the International Brain Initiative projects in a special 2019 neuroethics-focused issue of Neuron (https://www.cell.com/neuron/fulltext/S0896-6273(19)30068-6). While these are the first published set of neuroethics questions for neuroscientists, we anticipate that they can be further refined or expanded based on each project’s needs. BRAIN may also choose to adapt these further after deeper exploration of them in the second half of BRAIN.

### Table 4. Neuroethics Questions for Neuroscientists (NeQN)

1. What is the potential impact of a model or neuroscientific account of disease on individuals, communities, and society?
   1a. Possible unintended consequences on social stigma and self-stigma
   1b. Possible social or cultural biases in research design or interpretation of scientific results?

2. What are the ethical standards of biological material and data collection and how do local standards compare to those of global collaborators?
   2a. Protecting the privacy of human brain data (e.g. Images, neural recordings, etc.) and data, in immediate or legacy use beyond the experiment?
   2b. Special regard for brain tissue and its donors due to tissue origin and its past

3. What is the moral significance of neural systems that are under development in neuroscience research laboratories?
   3a. What requisite or minimum features of engineered neural circuitry generate concern about moral significance?
   3b. Are ethical standards for research adequate and appropriate for evolving methodologies and brain models?

4. How could brain interventions impact or reduce autonomy?
   4a. Identifying measures to ensure optimal autonomy and agency for participants/users
   4b. Responsibility for effects (where responsibility broadly encompasses legal, economic, and social contexts)

5. In which contexts might a neuroscientific technology/innovation be used or deployed?
   5a. Identifying applications that might be considered misuse or best uses beyond the laboratory?
   5b. Does this research raise different and unique equity concerns and, if so, have equitable access and stakeholder benefit been considered?

### BRAIN Neuroethics Working Group Guiding Principles

The NIH BRAIN Neuroethics Working Group proposed eight neuroethics guiding principles as points to consider for researchers, Institutional Review Boards (IRBs), and others involved in the conduct of BRAIN Initiative-funded research (Table 5). Two overarching principles frame the eight Neuroethics Guiding Principles (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6297371/): i) pursuing neuroscience research is an ethical imperative because of the immense suffering and economic impact of brain disorders around the world; and ii) neuroethics is vital to and should be integrated with neuroscience research. The Neuroethics Guiding Principles are meant to guide neuroscientists, particularly BRAIN Initiative-supported researchers, to help them consider the ethical, legal, and societal implications of their work in dialogue with other key stakeholders.

### Table 5. Neuroethics Guiding Principles

<table>
<thead>
<tr>
<th>Neuroethics Working Group, NIH BRAIN Initiative 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle</td>
</tr>
</tbody>
</table>

### Principle

<table>
<thead>
<tr>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Make assessing safety paramount</strong></td>
</tr>
<tr>
<td>Gene editing technologies such as CRISPR/Cas9 may offer hope for mitigating or eliminating brain disorders, yet risks and long-term effects are unknown, and there is potential for off-target effects. When researching innovative approaches, attend to preclinical data, monitor safety throughout, and inform participants about possible unexpected safety issues.</td>
</tr>
<tr>
<td><strong>Anticipate special issues related to capacity, autonomy, and agency</strong></td>
</tr>
<tr>
<td>Anticipate possible changes in preferences and agency, such as personality changes reported by some after deep brain stimulation for movement disorders; or deciding about control over stimulation parameters when brain stimulation paradigms target reward processing and motivation circuits. Seeking informed consent from participants, while simultaneously manipulating neural processes necessary for consent capacity and autonomous choice.</td>
</tr>
<tr>
<td><strong>Protect the privacy and confidentiality of neural data</strong></td>
</tr>
<tr>
<td>Protecting large, shared databases containing brain imaging data, as someday a brain MRI might be as identifying as a fingerprint. Determining who has access to personally identifiable information.</td>
</tr>
<tr>
<td><strong>Attend to possible malign uses of neuroscience tools and neurotechnologies</strong></td>
</tr>
<tr>
<td>Researchers’ have a responsibility to try to predict plausible misuses, prevent it when possible through design and security measures and ensure that participants, IRBs, government officials and others understand possible risks</td>
</tr>
<tr>
<td><strong>Use caution when moving neuroscience tools and neurotechnologies into medical or non-medical uses</strong></td>
</tr>
<tr>
<td>Discourage the premature widespread use or inappropriate adoption of new technologies such as neural markers of pain or deception, especially those offered directly to consumers or in non-health care settings, such as the legal system.</td>
</tr>
<tr>
<td><strong>Identify and address specific concerns of the public about the brain</strong></td>
</tr>
<tr>
<td>The public may worry that a beneficial improvement in ability to control a dysfunctional brain (e.g. from memory loss or seizures) has a flip-side, potentially threatening cognitive liberty. Or have justified concerns that research could “make a person someone else,” or result in entities that have morally significant human-like features.</td>
</tr>
<tr>
<td><strong>Encourage public education and dialogue</strong></td>
</tr>
<tr>
<td>Balancing appropriate understanding of neurological advances while avoiding hyperbole and correcting overly optimistic interpretations.</td>
</tr>
<tr>
<td><strong>Behave justly and share the benefits of neuroscience research and resulting technologies</strong></td>
</tr>
<tr>
<td>Identifying strategies to ensure wide sharing of the benefits of novel technologies and interventions and avoid exacerbating existing health disparities or inequalities.</td>
</tr>
</tbody>
</table>

Several other important efforts are not mentioned in detail in this Neuroethics Roadmap but have been developing scientific guidance for particular neurotechnologies and contexts for their use. These include the Human Performance Enhancement Report (https://www.amacad.org/news/regulatory-and-ethical-dimensions-human-performance-enhancement) from the American Academy.
of Arts and Sciences and Principles for Responsible Innovation in Neurotechnology led by policy organizations such as the Organization for Economic Cooperation and Development. This latter group has had a recent focus on neuroethics (http://www.oecd.org/sti/emerging-tech/workshop-on-minding-neurotechnology.htm) and how neuroethics might be integrated into public-private-partner led neuroscience research.

Ethical attention guided by frameworks such as those listed here, accompanied by careful reflection will continue to be essential when decisions are made about how to obtain knowledge about the brain and how to interpret it; about who uses the knowledge generated; as well as the implications of such knowledge for clinical practice, public health, other social institutions, and society. The selected frameworks, principles, and recommendations highlighted in this chapter provide guidance at multiple levels, for those who conduct, fund, disseminate, implement, and use neuroscience research.

In addition to the frameworks and principles described herein, important legal and regulatory requirements must be considered – for example, those that regulate the protection of humans and non-human animals in research (see Chapter 2. Understanding Ourselves: The Uniqueness of Neuroscience and Chapter 4: Neuroethics and Research on Animals, respectively). Scientists testing and developing emerging technologies should consider relevant general principles from the Belmont Report as well as the Presidential Commission’s principles for assessing emerging technologies. The BRAIN Initiative’s Neuroethics Guiding Principles, the Nuffield Council’s ethical framework, the NeQNs from the Global Neuroethics Summit, and issues and ideas that may be developed, are more targeted and specific in bringing attention to particular issues that arise in neuroscience and neurotechnology research. These latter resources raise many considerations relevant to BRAIN Initiative research, including the possible effects of neurotechnologies on agency, identity, capacity, and public trust, and risks associated with augmentation, hype, bias, and possible misuse of technologies and data. As brain research develops, it is possible – perhaps likely – that new concerns will arise that require additional consideration and that may point to refining guidelines or developing new ones. Going forward, the BRAIN Initiative should be prepared to support these discussions.

Integrating neuroethics and neuroscience

The BRAIN Initiative has emphasized the value of integrating neuroscience and cognitive science with technology and engineering, as well as encouraging neuroscience to be a boldly multidisciplinary exercise. To continue to confront challenging and emerging ethical questions arising from studying the brain, neuroethics benefits from integration with neuroscience – intentionally including scholarship from philosophy, psychology, law, theology, sociology, and other areas. Integrating a neuroethics perspective into neuroscience research design and conduct will have a powerful, positive impact on research and the knowledge it generates. Neuroethicists can help to scan the horizon and assist in anticipating and navigating ethical concerns, and they can also help guide how neuroscience research is designed, conducted, interpreted, and applied. Neuroethics should be intentionally integrated into neuroscience projects, but neuroethics research should also continue as independent scholarship that complements experimental neuroscience. Opportunities are many and include: i) seeking the advice of a neuroethicist on experimental design and details of research protocols; ii) collaborating with a neuroethicist to explore a unique ethical concern related to the implementation of an experiment or possible implications of study findings; or iii) collaborating with a neuroethicist to conduct parallel neuroethics research. Neuroethics research might be conceptual, normative, empirical, policy-related, or some combination of these (see text box below).

Types of Neuroethics Research

Conceptual (normative) neuroethics research

- Analysis of specific concepts such as privacy or personal identity
- Philosophical research about normative questions (i.e., what ought to constitute desirable or acceptable social behaviors?)
- Examples: How should one define and treat people with various levels of consciousness? Does fluctuating capacity from disease, a brain injury, or a brain intervention indicate a need to rethink informed consent?

Conceptual and normative neuroethics research may draw from existing literature and theories, as well as practices from law, philosophy, theology, and neuroscience. The October 2018 issue of the journal Neuroethics, for example, published conceptual papers on neuroscience and free will, self-governance, self-control, and decision-making.

Empirical neuroethics research
Systematic data collection to ascertain views, values, or practices of researchers, patients, research participants, or the public

May employ social-science methodologies such as quantitative surveys or qualitative interviews and could also include experimental designs to test the impact of interventions or other experimental manipulations.

Examples: The BRAIN Initiative has funded a number of neuroethics empirical projects.

The need for neuroethics research

Like the Human Genome Project, the BRAIN Initiative is a wide-ranging endeavor that can raise equally wide-ranging ethical, social, and legal issues. Tools and knowledge emanating from decoding the human genome transformed biomedicine dramatically. In the decades hence, individual labs across the globe—not to mention citizen scientists and children in school—have ready access to relatively easy-to-use methods to “read” DNA that have found ways to bypass regulatory scrutiny (https://www.nature.com/articles/nbt.3761). Such access has opened many new doors of investigation, launching numerous new fields of ’omics inquiry as well as numerous controversies. Newfound experimental access to our genome has even coined many phrases, such as the “language of life.”

Another important consequence of the Human Genome Project has been a tendency for people in many segments of society to embrace a form of genetic essentialism—some people equate “who we are” with our genes. Will the BRAIN Initiative have a similar reductionist effect on how we as humans view ourselves? Will we equate “who we are” with brain function at the expense of alternative, more relational conceptions of identity? Is society prepared?

These are not just rhetorical questions, but a call for systematic neuroethics research to learn how neuroscience will have impacts beyond the bench and how it will set new societal norms. Moreover, unlike blood samples and DNA, brain tissue and the data derived from it might be considered more sensitive given the connection between the brain and personally identifying behavior. Well-known ethical standards and practices for human research participants are in place—including institutional requirements for IRB review, data safety and monitoring, and other formal oversight mechanisms. Because modern neuroscience will continue to pose difficult ethical challenges for research with human participants, relying upon existing biomedical research ethics guidelines alone will not be adequate to contend with the aims and consequences of the BRAIN Initiative.

In summary, neuroethics is integral to the BRAIN Initiative and cannot be separated from it. Neuroethics provides an opportunity for deliberation, analysis, and research that both catalyzes, improves, and enables neuroscience. This Neuroethics Roadmap charts the way forward to maximize innovation and value from the BRAIN Initiative in a way that prioritizes benefits for humanity at large. To do so, it explains what neuroethics can offer, provides neuroethics principles and guidelines to help shape ethical neuroscience and its applications, promotes neuroethics research, and endorses integration of neuroethics with neuroscience at multiple levels, through the articulation of implementable goals.

CHAPTER 2. STUDYING OURSELVES: THE UNIQUENESS OF NEUROSCIENCE

CHAPTER 3. NEUROETHICAL ISSUES AND NEUROTECHNOLOGIES

CHAPTER 4. NEUROETHICS AND RESEARCH WITH ANIMALS

CHAPTER 5. BEYOND THE BENCH: REAL-WORLD TRANSLATION OF NEUROSCIENCE RESEARCH

CHAPTER 6. INTEGRATING NEUROETHICS AND NEUROSCIENCE

NEUROETHICS MOONSHOT - Revolutionizing BRAIN: The Theory of the Mind
Priority Areas
Funding Opportunities

Cell Type (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=1)

Circuit Diagrams (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=18)

Monitor Neural Activity (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=2)

Interventional Tools (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=3)

Theory & Data Analysis Tools (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=4)

Human Neuroscience (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=5)

Integrated Approaches (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=19)

BRAIN Update (https://www.brainupdate.nih.gov/)

BRAIN 2025 Report (/strategic-planning/brain-2025-report)
Neuroethics Roadmap | Brain Initiative

(about/brain-social-media) (about/brain-social-media) (about/brain-social-media)

GRAY MATTERS
Topics at the Intersection of Neuroscience, Ethics, and Society

Presidential Commission for the Study of Bioethical Issues

March 2015
On the cover: Illustrated representation of the surface of the brain and spinal cord.
GRAY MATTERS
Topics at the Intersection of Neuroscience, Ethics, and Society

Presidential Commission
*for the* Study of Bioethical Issues

Washington, D.C.
March 2015

http://www.bioethics.gov
Dr. John Arras (1945-2015) was a consummately dedicated teacher, lauded moral philosopher, and an eminent scholar of bioethics. He brought out the very best in everyone who had the privilege and pleasure of working with and learning from him. For the past five years, we were honored to have John as a thoroughly engaged and beloved member of our Presidential Commission for the Study of Bioethical Issues. In the words of Commission member Stephen Hauser, John was an “irreplaceable member” of our group. We have lost, as Commission member Nelson Michael wrote, “a dear friend, colleague, and one of the greats of bioethics.”

John contributed far more than his share to our Commission’s painstaking work. He had an unparalleled gift for bringing philosophical insight to thorny medical and scientific conundrums. Even that gift paled in comparison to John’s wry, perfectly timed humor. Due in no small part to his flair for intellectual provocation—as feisty as it was friendly—our Commission rapidly became, as Vice Chair James Wagner keenly observed, something more than a commission. We became a fondly argumentative and loving extended family with John, as Commission member Raju Kucherlapati said, “the lightning rod for many discussions.” Commission member Barbara Atkinson captured John’s quintessential character as “one of the most thoughtful and giving people I have known. He was strong in his views but open to discussion and compromise, so he was extremely valuable for our discussions and final reports.”

As a lover of learning and seeker of justice for all, John Arras was as good as we can ever hope to get. We shall carry forth John’s spirit as we grieve the tremendous loss of a great teacher, scholar, and member of our bioethics family. We already miss him dearly.

Dr. Amy Gutmann
Chair, Presidential Commission for the Study of Bioethical Issues
March 9, 2015
ABOUT THE PRESIDENTIAL COMMISSION FOR
THE STUDY OF BIOETHICAL ISSUES

The Presidential Commission for the Study of Bioethical Issues (Bioethics Commission) is an advisory panel of the nation’s leaders in medicine, science, ethics, religion, law, and engineering. The Bioethics Commission advises the President on bioethical issues arising from advances in biomedicine and related areas of science and technology. The Bioethics Commission seeks to identify and promote policies and practices that ensure scientific research, health care delivery, and technological innovation are conducted in a socially and ethically responsible manner.

For more information about the Bioethics Commission, please see http://www.bioethics.gov.

The use of trade names and commercial sources in this report is for identification only and does not imply endorsement.
Dear Mr. President:

On behalf of the Presidential Commission for the Study of Bioethical Issues, we present to you *Gray Matters: Topics at the Intersection of Neuroscience, Ethics, and Society*, the second part of the Bioethics Commission’s response to your request of July 1, 2013. In its first volume, *Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society*, the Bioethics Commission analyzed why and how to achieve ethics integration early and explicitly throughout neuroscience research. In this second and final volume, the Bioethics Commission broadly considered the ethical and societal implications of neuroscience research and its applications.

Building on its earlier work, the Bioethics Commission addressed this topic in nine public meetings, where it heard from experts from myriad disciplines and perspectives, including neuroscientists, philosophers, educators, ethicists, federal regulators, public- and private-sector partners involved in the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, and representatives of affected communities with a stake in the outcomes of neuroscience research. In addition, the Bioethics Commission solicited public comment and received more than 30 thoughtful responses.

Contemporary neuroscience offers the opportunity to better understand the human brain and support the development of more effective diagnostic tools, treatments, preventions, and cures for neurological disorders and psychiatric conditions that affect tens of millions of individuals in the United States alone.
This promise—along with the potential to gain a deeper understanding of our cognition, emotion, imagination, behavior, memory, learning, and social interactions—has captured the interest of scientists and the public alike. The Bioethics Commission delved deeply into three important topics that advancing neuroscience and technology throw into heightened ethical and practical relief: cognitive enhancement, consent capacity, and neuroscience and the legal system.

This report seeks to clarify for the public the current scientific landscape, clear a path to productive discourse to navigate difficult issues as they arise, and identify common ground where it exists. We offer 14 recommendations to guide the ethical progress of neuroscience research and its applications. Our recommendations call for attention to fundamental ethical concerns regarding, for example, justice and stigmatization of groups and individuals; research to clarify persistent questions and fill gaps in our current state of knowledge; accurate communication about the ethical and practical implications and application of neuroscience research results; clarity around legal requirements and new guidance where needed; and the need to support and advance innovative multidisciplinary research and scholarship at the critically important intersection of neuroscience, ethics, and society.

The Bioethics Commission is honored by the trust you have placed in us, and we are grateful for the opportunity to serve you and the nation in this way.

Sincerely,

Amy Gutmann, Ph.D.  
Chair

James W. Wagner, Ph.D.  
Vice Chair
The White House
Washington, D.C.
July 1, 2013

The Honorable Amy Gutmann, Ph.D.
Commission Chair
Presidential Commission for the Study of Bioethical Issues
Washington, D.C. 20005

Dear Dr. Gutmann:

As I noted in my announcement of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative on April 2, 2013, developments in neuroscience hold great potential to help individuals and society. New technologies to better visualize the brain and understand how it works promise to speed the discovery of new ways to treat and prevent brain disorders, including those caused by disease and traumatic injury, and to shed light on the neural components of memory and learning, among other benefits.

Advances in neuroscience can also raise ethical and legal issues that require reflection and analysis. In keeping with my Administration’s strong commitment to rigorous research ethics in all fields, I want to ensure that researchers maintain the highest ethical standards as the field of neuroscience continues to progress. As part of this commitment, we must ensure that neuroscientific investigational methods, technologies, and protocols are consistent with sound ethical principles and practices.

Equally important, we should consider the potential implications of the discoveries that we expect will flow from studies of the brain, and some of the questions that may be raised by those findings and their applications—questions, for example, relating to privacy, personal agency, and moral responsibility for one’s actions; questions about stigmatization and discrimination based on neurological measures of intelligence or other traits; and questions about the appropriate use of neuroscience in the criminal-justice system, among others. It will also be important to consider these types of questions as they relate to different life stages, from infancy through old age.
I request that the Presidential Commission for the Study of Bioethical Issues engage with the scientific community and other stakeholders, including the general public, to identify proactively a set of core ethical standards—both to guide neuroscience research and to address some of the ethical dilemmas that may be raised by the application of neuroscience research findings.

In the course of your deliberations, I encourage you to reach out to a wide range of constituencies, including scientists, ethicists, legal scholars, and members of the public, to ensure that your findings and the neuroscience enterprise faithfully reflect and strengthen our values as a Nation.

Sincerely,
PRESIDENTIAL COMMISSION FOR THE STUDY OF BIOETHICAL ISSUES

STAFF* AND CONSULTANTS

Executive Director
Lisa M. Lee, Ph.D., M.A., M.S.

Associate Directors
Michelle Groman, J.D.
Kaye Spector-Bagdady, J.D., M. Bioethics

Communications Director
Hillary Wicai Viers, M.S.J.

Senior Advisors
Paul A. Lombardo, Ph.D., J.D.
Jonathan D. Moreno, Ph.D.

Research Staff
Misti Ault Anderson, M.S., M.A.
Rachel S. Bressler, J.D.
Kata Chillag, Ph.D.
Elizabeth M. Fenton, Ph.D., M.P.H.
Karen M. Meagher, Ph.D.
Olivia Nevitt, M.P.H.
Cristina Nigro, M.S.
Elizabeth R. Pike, J.D., LL.M.
Maneesha Sakhuja, M.H.S.
Michelle Spektor, B.S.
Nicolle K. Strand, J.D., M. Bioethics
Michael N. Tennison, M.A.
Victoria Wilbur, B.A.
Tenny R. Zhang, B.A.

Consultants
Burness Communications
C. Kay Smith, M.Ed.
Medical Arts Branch, Office of Research Services, National Institutes of Health

Communications Staff
Alannah Kittle, M.P.H.

Administrative Staff
Tynetta Dreher
Esther E. Yoo

Interns
Ruqayyah Abdul-Karim, B.A.
Cassandra Hunter, M.A.
Minjung Koo, B.A.
Rahul Nayak, B.S.
Abbas Rattani, M. Bioethics
Hillary Samples, M.H.S.
Haley Sullivan
Kevin Tobia, B.A.
Sean Valle, B.A.
Tiana Woolridge
Abena Yeboa, B.S., B.A.

*Includes former research staff
ACKNOWLEDGEMENTS

The Bioethics Commission is grateful to the many people who contributed their time, effort, and expertise to this report. Over 60 diverse and distinguished speakers participated in our public meetings, engaged in thought-provoking discussions, and guided our recommendations. We thank them for their service and willingness to engage in the deliberative process.

The Bioethics Commission also thanks our dedicated and invaluable staff who provided thoughtful guidance, thorough research, and steadfast support throughout our deliberations on ethics and neuroscience. The Bioethics Commission extends special thanks to Executive Director Lisa M. Lee for her committed and thoughtful leadership on all of our work, Associate Director Michelle Groman for her adept direction and painstaking attention to detail, and Senior Advisor Jonathan D. Moreno for his deft advice and contributions to this report. The Bioethics Commission is especially appreciative to staff lead Misti Ault Anderson for her unwavering support and perseverance in facilitating our deliberation of this considerable and meaningful topic. Finally, we thank Kata Chillag and Nicolle K. Strand for their tireless efforts in support of the drafting of both volumes of this report.
EXECUTIVE SUMMARY
Neuroscience presents an unparalleled opportunity to gain a deeper understanding of the human brain and mind, including our cognition, behavior, memory, learning, mood, and social interactions. It also offers new opportunities to treat, prevent, and possibly cure neurological disorders that constitute an immense public health burden worldwide.

In 2013, President Obama announced the federal Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, and charged the Presidential Commission for the Study of Bioethical Issues (Bioethics Commission) to examine ethical considerations related to both the conduct of neuroscience research and the application of neuroscience research findings. The Bioethics Commission addressed the President’s charge in two parts. In its first volume on neuroscience and ethics, *Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society*, the Bioethics Commission emphasized the importance of integrating ethics and neuroscience throughout the research endeavor. This second volume, *Gray Matters: Topics at the Intersection of Neuroscience, Ethics, and Society*, takes an in-depth look at three topics at the intersection of neuroscience and society that have captured the public’s attention.

The Bioethics Commission found widespread agreement that contemporary neuroscience holds great promise for relieving human suffering from a number of devastating neurological disorders. Less agreement exists on multiple other topics, and the Bioethics Commission focused on three cauldrons of controversy—cognitive enhancement, consent capacity, and neuroscience and the legal system. These topics illustrate the ethical tensions and societal implications of advancing neuroscience and technology, and bring into heightened relief many important ethical considerations.

The Bioethics Commission seeks to clarify for the public the scientific landscape, identify common ground for productive discourse surrounding these topics, and recommend an ethical path forward to support the progress of neuroscience research.

**Cognitive Enhancement and Beyond**

In this report, the Bioethics Commission expanded the conversation beyond the ongoing cognitive enhancement debate to include other forms of neural modification. First, it considered not only novel neurotechnologies, but also
methods, behaviors, and conditions that alter the brain and nervous system. Second, it expanded the discussion beyond use of products and methods that enhance cognition to include those that alter the brain or nervous system in a wide range of ways, such as altering motor function.

Neural modification can serve at least three purposes, to (1) maintain or improve neural health and cognitive function within typical or statistically normal ranges; (2) treat disease, deficiency, injury, impairment, or disorder (referred to as “neurological disorders”) to achieve or restore typical or statistically normal functioning; and (3) expand or augment function above typical or statistically normal ranges. In delineating these neural modification objectives, the Bioethics Commission is mindful that they are not always sharply distinguishable.

Alternating the brain and nervous system is not inherently ethical or unethical. Ethical assessment of neural modification requires consideration of who is choosing the modifier, what is being chosen, what its purposes are, who stands to benefit, and who might be harmed. Members of the public must be well-informed to make educated, practical decisions about personal health and wellbeing, and participate in collective deliberation and decision making about societal applications of neural modifiers.

Several well-known lifestyle interventions, such as adequate sleep, exercise, and nutrition, are associated with improved neural function. Similarly, public health interventions, such as lead paint abatement, can help prevent the negative impact of environmental exposures on neural development and function. These behaviors and conditions can maintain and improve neural health and might be safer and more effective than those offered by novel neural modifiers, some of which might have minimal benefits and uncertain risks.

**Recommendation 1: Prioritize Existing Strategies to Maintain and Improve Neural Health**

In addition to developing new drugs and devices to maintain and improve neural health, funders should prioritize and support research on existing, low-technology strategies, such as healthy diet, adequate exercise and sleep, lead paint abatement, high-quality educational opportunities, and toxin-free workplaces and housing.
Existing treatments for neurological disorders are valuable and can be improved. Emerging neural modification interventions will help reduce the individual and societal burden of neurological disorders. Safe and effective treatments can improve the lives of millions of individuals living with such conditions.

**Recommendation 2: Prioritize Treatment of Neurological Disorders**

Funders should prioritize research to treat neurological disorders to improve health and alleviate suffering. This research should consider individual, familial, and public health burdens as well as potential risks, benefits, and long-term effects of specific interventions.

Although the Bioethics Commission recognizes the need to prioritize the study of both traditional and novel interventions for the prevention and treatment of neurologic disorders, it nonetheless also supports research to better characterize and understand novel neural modification techniques to augment or enhance neural function. Limited, inconclusive evidence exists for the benefits and risks of stimulant drugs and brain stimulation methods as neural enhancers. In addition, few data are available on the prevalence of the use of neural modification interventions for cognitive enhancement purposes.

**Recommendation 3: Study Novel Neural Modifiers to Augment or Enhance Neural Function**

Funders should support research on the prevalence, benefits, and risks of novel neural modifiers to guide the ethical use of interventions to augment or enhance neural function.

If safe and effective novel forms of cognitive enhancement become available, they will present an opportunity to insist on a distribution that is fair and just. While not eliminating all other less tractable forms of injustice in the distribution of neural health and wellbeing, it is possible to ensure that any new forms of safe and beneficial neural modification do not worsen those injustices.

Limiting access to effective enhancement interventions to those who already enjoy greater access to other social goods would be unjust. It also might deprive society of other benefits of more widespread enhancement that increase as more individuals have access to the intervention. In addition, more widespread enhancement might help to close some gaps in opportunity that are related to neural function, such as educational attainment or employment.
Recommendation 4: Ensure Equitable Access to Novel Neural Modifiers to Augment or Enhance Neural Function

Policymakers and other stakeholders should ensure that access to beneficial, safe, effective, and morally acceptable novel neural modifiers to augment or enhance neural function is equitable so as not to compound or exacerbate social and economic inequities.

Clinicians often receive requests to prescribe medications for cognitive enhancement, and they must decide whether to prescribe the medication to particular patients. These decisions are more ethically complex with regard to children, because children lack legal and ethical consent capacity and are vulnerable to coercion. Other stakeholders also would benefit from education and guidance on neural modification interventions. These stakeholders include employers, parents, educators, and professional organizations in fields such as aviation, medicine, and the military, among others, that are associated with on-the-job use of brain and nervous system enhancement interventions.

Recommendation 5: Create Guidance About the Use of Neural Modifiers

Professional organizations and other expert groups should develop guidance for clinicians, employers, parents, educators, and patients about the use of neural modifiers and their potential risks and benefits. Medical professional organizations should develop guidelines to assist clinicians in responding to requests for prescriptions for interventions to expand or augment neural function. Clinicians should not prescribe medications that have uncertain or unproven benefits and risks to augment neural function in children and adolescents who do not have neurological disorders.

Capacity and the Consent Process

Neuroscientists who conduct research involving human participants commonly work with populations or individuals whose consent capacity might be absent, impaired, fluctuating, or in question. Understanding, evaluating, and improving informed consent processes is an ongoing goal of research ethics. The history of research ethics includes multiple efforts by national-level advisory bodies to provide guidance for research involving individuals who might have compromised or impaired consent capacity. This history illustrates the challenging tension between the need for rigorous research on important
diseases and conditions, and the need to protect individuals who might be vulnerable because of impaired consent capacity. Federal regulations specific to research involving individuals with impaired consent capacity have never been adopted.

Neuroscience research is an important means of promoting progress and benefiting populations affected by neurological disorders and psychiatric conditions, including those associated with impaired consent capacity, and should proceed with adequate ethical safeguards and protections in place. This dual mission—protection and inclusion to ensure that the benefits of research are distributed equitably—shapes many core ethical considerations surrounding capacity, the consent process, and participation in research.

**Recommendation 6: Responsibly Include Participants with Impaired Consent Capacity in Neuroscience Research**

Researchers should responsibly include individuals with impaired consent capacity who stand to benefit from neuroscience research. Participation, with ethical safeguards in place, can ensure progress aimed at understanding and ameliorating neurological disorders and psychiatric conditions.

Researchers have made substantial progress in the past decade in characterizing and understanding consent capacity. However, gaps remain, and further research can support development of best practices for ethical research involving participants with impaired consent capacity. Conceptual research on gaps in our knowledge, including the influence of vulnerability, desperation, and affective states on decision making, could lead to better protections for all research participants. Moreover, empirical research evaluating assessment tools and additional protections for participants with impaired consent capacity can determine whether they are adequately protective.

**Recommendation 7: Support Research on Consent Capacity and Ethical Protections**

Funders should support research to address knowledge gaps about impaired consent capacity, including the concept of capacity, brain function and decision-making capacity, current policies and practices, and assessment tools.
Equating certain conditions with impaired consent capacity or making unfounded assumptions about individual abilities based on diagnoses can exacerbate or perpetuate stigma. Ethical neuroscience research can foster a more accurate understanding of neurological disorders and mental illness, and potentially mitigate stigma. One principal approach to help neuroscience researchers alleviate stigma is stakeholder engagement.

**Recommendation 8: Engage Stakeholders to Address Stigma Associated with Impaired Consent Capacity**

Funders and researchers should engage stakeholders, including members of affected communities, to build understanding of consent capacity and associated diagnoses to mitigate the potential for stigma and discrimination.

Including affected individuals (those with impaired consent capacity and others) in research is vital to fulfill the promise of neuroscience to ameliorate neurological disorders and psychiatric conditions. The Common Rule—the regulations supplying standards for the ethical conduct of federally supported human subjects research—requires informed consent from research participants or legally authorized representatives (LARs) before research can proceed. Thus, an important step in conducting ethical research involving individuals with impaired consent capacity is determining who can serve as an LAR. Federal and state laws lack clarity about how to make this determination.

**Recommendation 9: Establish Clear Requirements for Identifying Legally Authorized Representatives for Research Participation**

State legislatures and federal regulatory bodies should establish clear requirements to identify who can serve as legally authorized representatives for individuals with impaired consent capacity to support their responsible inclusion in research.

**Neuroscience and the Legal System**

Advances in neuroscience offer a better understanding of human behavior, the potential for improved policymaking, increased accuracy, and decreased errors in advancing justice. The application of neuroscience to the law also raises concerns about scientific reliability, misapplication and overreliance on a developing science, conceptions of free will, mental privacy, and personal
liberty. Although neuroscience might help us achieve more accuracy in decision making and better policies for trials and sentencing, it does not change the normative or moral questions that the law seeks to answer. Law is a social institution, built on norms developed and instituted by society.

Neuroscience has multiple potential applications to the legal system and already is employed in many relevant contexts. Prosecutors and defense attorneys use neuroscience evidence in criminal proceedings to support propositions concerning, for example, competency to stand trial, mitigation of criminal responsibility, and predicting future dangerousness. Parties also use neuroscience evidence in the civil context to provide objective evidence of “invisible” injuries such as toxic exposure, pain, and suffering. Policymakers have invoked neuroscience to advocate for legislation and reform; scholars have advocated for use of neuroscience to address biases in legal decision making; and researchers and even some commercial entities have introduced novel uses of neuroscience for investigative purposes.

Members of the public, especially ones who will serve as jurors, can benefit immensely from educational resources that help bring high-level neuroscientific concepts into lay terms. Individuals expected to use and interpret neuroscience, including judges and attorneys, can also benefit from greater availability of basic training that helps ease the interdisciplinary transition of neuroscience into the legal decision-making process and effectively assess the evidence and technologies involved in a growing number of legal cases.

**Recommendation 10: Expand and Promote Educational Tools to Aid Understanding and Use of Neuroscience within the Legal System**

Government bodies and professional organizations, including legal societies and nonprofit organizations, should develop, expand, and promote training resources, primers, and other educational tools that explain the application of neuroscience to the legal system for distribution to members of the public, jurors, judges, attorneys, and others.

In addition to the broad educational tools discussed in Recommendation 10, relevant bodies also should fund and conduct specific research and report results regarding use of neuroscience evidence in making important legal and policy decisions. Organizations and government bodies also should publish
reports that address the challenges and limitations of neuroscience’s application to the legal system.

Recommendation 11: Fund Research on the Intersection of Neuroscience and the Legal System

Relevant bodies, such as the National Academies of Science, the U.S. Department of Justice, the National Institute of Justice, and the Social Security Administration, should support comprehensive studies of the use of neuroscience in legal decision making and policy development.

Neuroscience can add value to legal decision making and policy development. To maximize the value that neuroscience has to offer, scientists, the media, and legal decision makers must avoid hype. When neuroscience evidence that is unreliable or has not yet been validated and is not ready for application is introduced into the legal system, justice is threatened. Unrealistically high expectations for new science and technology can lead to a loss of trust when those expectations are unmet.¹⁰

Recommendation 12: Avoid Hype, Overstatement, and Unfounded Conclusions

Neuroscientists, attorneys, judges, and members of the media should not overstate or rely too heavily on equivocal neuroscientific evidence to draw conclusions about behavior, motivations, intentions, or legal inferences.

As attorneys introduce more neuroscience evidence into the courtroom, and advocates use neuroscience to influence policy, neuroscientists should engage with the process, consider potential legal applications of their work, and seek to engage with legal and policy decision makers to ease the translation. Neuroscientists can play a principal role in assisting judges and jurors with determining the appropriate interpretations of neuroscientific evidence.¹¹

Recommendation 13: Participate in Legal Decision-Making Processes and Policy Development

Neuroscientists should participate in legal decision-making processes and policy development to ensure the accurate interpretation and communication of neuroscience information.
Conclusion

In this report, the Bioethics Commission calls for research on a number of critical topics. Such research requires adequate support, including funding, personnel, and other resources. As a White House Grand Challenge, the BRAIN Initiative is uniquely positioned to establish and support efforts that bring together diverse expertise from neuroscience, ethics, law, policy, and other disciplines to advance research and education at the intersection of neuroscience, ethics, and society.

Recommendation 14: Establish and Fund Multidisciplinary Efforts to Support Neuroscience and Ethics Research and Education

The BRAIN Initiative should establish and fund organized, independent, multidisciplinary efforts to support neuroscience and ethics research and education, including the activities recommended in this report.

***

Neuroscience advances have captured the public’s attention and stimulated scholarly and public debate, fueled by accurate accounts of the science as well as hyped or misinformed interpretations. Three controversies are some of the most important and provocative topics at the intersection of neuroscience and society. Neural modification, including cognitive enhancement, raises questions about reconciling risks and benefits, ensuring justice, and understanding what it means to be human. Adequately respecting and protecting individuals with impaired consent capacity has presented challenges for decades. Advances in neuroscience and the promise of neuroscience research compel us to reexamine this area—ensuring that those with impaired consent capacity can participate in and benefit from ethical research. Application of neuroscience to legal decision making and policy development offers potential for more accurate and just outcomes, but also raises concerns about premature use of scientific information, privacy, and moral responsibility. In this report, the Bioethics Commission seeks to clarify the scientific landscape, identify common ground, and recommend ethical paths forward to stimulate and continue critical, well-informed conversations at the intersection of neuroscience and ethics as the field continues to advance.
CHAPTER 1
Introduction
Advances in contemporary neuroscience research offer the prospect of great individual and societal benefit. Neuroscience presents an unparalleled opportunity to gain a deeper understanding of the human brain and mind, including our cognition, behavior, memory, learning, mood, and social interactions. It also offers new opportunities to treat, prevent, and possibly cure neurological disorders that constitute an immense public health burden worldwide. In 2013, President Obama announced the federal Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, harnessing a diverse array of experts to pursue discoveries that will have a “lasting positive impact on lives, the economy, and our national security.”

The President charged the Presidential Commission for the Study of Bioethical Issues (Bioethics Commission) to examine ethical considerations related to both the conduct of neuroscience research and the application of neuroscience research findings.

The Bioethics Commission addressed the President’s charge in two parts. In its first volume on neuroscience and ethics, Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society (Gray Matters, Vol. 1), the Bioethics Commission emphasized the importance of integrating ethics and neuroscience throughout the research endeavor. This second volume, Gray Matters: Topics at the Intersection of Neuroscience, Ethics, and Society (Gray Matters, Vol. 2), takes an in-depth look at three particularly controversial topics at the intersection of neuroscience and society that have captured the public’s attention.

After engaging experts and reviewing public comments, the Bioethics Commission found widespread agreement that contemporary neuroscience holds great promise for relieving human suffering from numerous devastating neurological disorders. Less agreement exists on multiple other topics, and the Bioethics Commission resolved to focus its second Gray Matters volume on three cauldrons of controversy—cognitive enhancement, consent capacity, and neuroscience and the legal system. These three topics illustrate the ethical tensions and societal implications of advancing neuroscience and technology, and bring into heightened relief many important ethical considerations.
These controversial topics sit at the epicenter of both scholarly debate and public dialogue in popular media about the reach and impact of neuroscience. Alongside well-informed and well-intentioned voices, hyperbole and misinformation permeate the conversation—this exaggeration can lead to undue excitement and attention, commonly referred to as “hype.” Overstated interpretation of research results and expectations can mislead the public, misdirect resources, and instill misplaced fears. In this report, the Bioethics Commission aims to clear a path for productive discourse and the ethical progress of neuroscience research—progress that is crucial to alleviating serious medical conditions and bettering society in numerous other ways. Drawing from its collective expertise and crucial role as public educator, the Bioethics Commission offers this report to clarify for the public the scientific landscape, identify common ground, and recommend ethical paths forward on these topics.

**Cognitive Enhancement and Beyond**

In universities and colleges, doctors’ offices, secondary schools, professional societies, and popular newspapers and blogs, scholars, professionals, caregivers, and others actively debate the use of novel neuroscience products and other methods to enhance cognition. In this report, the Bioethics Commission clarifies and expands the conversation beyond the common cognitive enhancement debate. First, this report considers not only novel neurotechnologies, but also methods, behaviors, and conditions that alter the brain and nervous system. Second, the discussion moves beyond the use of products and methods that enhance *cognition* to include those that alter the brain or nervous system in wide-ranging ways, such as altering motor function.

The definition of enhancement and the distinction between enhancement and other forms of neural modification are at the center of this debate. For example, *The Neuroethics Blog* hosted by the Center for Ethics in the Neuroethics Program at Emory University (Atlanta, Georgia) regularly posts brief pieces covering topics at the intersection of neuroscience and ethics. A 2014 post outlined the difficulty in defining “normal” and distinguishing treatment and enhancement. Another post questioned whether diminishment of certain capabilities can be considered a form of enhancement.
Journalists also have reported on other ethical aspects of cognitive enhancement. For example, news stories have raised questions about both ethical and scientific controversy surrounding cognitive enhancement, including whether the wakefulness-promoting drug modafinil (Provigil®) boosts cognition, whether transcranial direct current stimulators can or should be regulated by the U.S. Food and Drug Administration (FDA), whether it is ethically permissible to erase bad memories, and whether brain augmentation is ethically justifiable.¹⁷

Scholars and the public also express concern over the use of direct-to-consumer (DTC) neuroscience products purported to enhance cognition. A 2014 consensus statement issued by a consortium of scientists rejects the claim that currently marketed brain games reduce or reverse cognitive decline, and several media outlets reported on the consensus.¹⁸ In addition, a report on one DTC neuroscience company’s announcement that it would allow researchers access to its vast repository of data collected from user play of its brain training games included skepticism about the accuracy of the data and possible inadequacy of privacy protections.¹⁹

Several stakeholders, including national and international leaders in neuroscience and ethics, urged the Bioethics Commission to consider cognitive enhancement and related topics. For example, the International Neuroethics Society considers human enhancement to be one of the most important ethical considerations related to neuroscience, focusing its comments to the Bioethics Commission on justice concerns, product safety and effectiveness, potential impact of DTC products, and implications for use of neural modifiers with healthy children and adolescents.²⁰ Against a background of strongly held and often opposing views, cognitive enhancement and neural modification are ripe for the Bioethics Commission to consider.
ETHICS INTEGRATION AND MITIGATING HYPE

Contemporary neuroscience offers great promise for attaining a deeper understanding of the human brain and alleviating the burden of neurological disorders, but overstatement and hype persist. Neuroscientists, members of the media who report on neuroscience advances, and others sometimes make claims that can be false, misleading, distorted, or exaggerated.

For example, the New York Times Magazine published an article in January 2015 about a researcher attempting to map all of the neural connections in the brain. The magazine’s cover depicts the words “This Is Your Brain” in letters formed by illustrated neurons. Prominent captions on the magazine’s cover and article cover page suggest that we could learn “everything” about ourselves through such a map, and ask, “If he succeeds, could we live forever as data?”

Less prominently, in its final paragraphs, the article presents critiques from scientists who argue that brain mapping would be “absolutely necessary but completely insufficient” to fully understand the human brain.

In Gray Matters, Vol. 1, the Bioethics Commission emphasized the importance of integrating ethics and neuroscience throughout the research endeavor, in part by integrating the perspectives of philosophers, social scientists, and others into the research process. An integrated perspective on the research described in this article might point out that even a complete map of the human brain would not fully inform our understanding of consciousness and the self. Articles that feature attention-grabbing headlines, but lack complete and balanced perspectives, serve to inform and excite the public about important neuroscience research, but run the risk of distorting and exaggerating its actual potential.


Capacity and the Consent Process

The burden of neurological disorders, mental illnesses, brain and nervous system injuries, and age-related cognitive decline is enormous. Progressing through human life without being affected by a decline in brain or nervous system function or needing to care for an affected loved one is likely impossible. Neuroscience offers great promise to reduce this burden and potentially find ways to prevent, treat, and cure many brain-related disorders and injuries. Clinical research involving affected populations is necessary to achieve this worthy goal. Managing a robust informed consent process with individuals who might have impaired consent capacity is especially pertinent in contemporary neuroscience. Impaired consent capacity is a hotly debated topic in research ethics, yet regulations, guidance, and research practices have remained essentially unchanged for decades. The Bioethics Commission is poised to move the conversation forward.
Although members of the public might not explicitly consider the topic of consent capacity, they are deeply affected by it. Even the healthy among us might one day benefit from interventions that preserve life or neurological function, borne from the fruits of neuroscience research that requires enrollment of individuals with impaired consent capacity. The popular media reported on some of the scholarly debate around this topic. For example, media sources reported on the FDA’s informed consent draft guidance, released in 2014, acknowledging that it addresses individuals “with diminished or fluctuating consent capacity.”

Others criticized the FDA’s position on legally authorized representatives in light of a dearth of state laws indicating who can legally serve as a representative in research. A 2014 post on The Neuroethics Blog considers the elements of ethical informed consent from individuals who might have impaired consent capacity in situations when the investigator stands to gain financially but the research participant will not benefit from the research.

Several stakeholders urged the Bioethics Commission to weigh in on the consent capacity discussion. For example, Public Responsibility in Medicine and Research (PRIM&R) identified numerous reports and recommendations from national bodies on consent capacity, inquired as to why many of those recommendations have not been implemented, and urged consistency in the field. Given the ongoing challenge to respect and adequately protect those with impaired consent capacity—who stand to gain the most, but who also would shoulder much of the burden of research—and the potential for contemporary neuroscience to help guide ethical policies, the Bioethics Commission addresses the topic in this report.

**Neuroscience and the Legal System**

The fruits of neuroscience research are applied in a variety of contexts, increasingly extending beyond the laboratory, from clinical translation to development of DTC products. Use of neuroscience in legal decision making and policy development is a controversial matter that has captured the public’s attention. Application of neuroscience to the legal system ranges from using evidence about the development of adolescent brains to prevent a teenager from receiving the death penalty, to using imaging technology to persuade a jury that a defendant might not be capable of the moral responsibility required for conviction of a crime.
Use of neuroscience within the legal system raises fundamental questions about how to determine when scientific findings are ready for public use, and how to ensure that scientific experts who testify and play a role in determining the fate of defendants are reliable and their conclusions sound. Such use of neuroscience raises questions about moral responsibility and biological determinism—whether our brains cause us to act badly and, if so, whether we can or should be held responsible for these actions.

The popular media have brought these questions to the public’s attention. The Public Broadcasting Service television series *Brains on Trial* features actor Alan Alda asking various experts if, and how, neuroscience research might be used in court, and whether scientific advances might challenge traditional notions of moral responsibility. Media articles include discussion about whether neuroscience can prove or disprove that we have free will, and how that might affect our notion of criminal responsibility. A 2014 article discusses how neuroscience evidence increasingly might play a role in decisions about whether one is competent to stand trial, or whether one should be spared a particular punishment because of a brain abnormality. NBC News reported on a new technique that uses brain waves to detect whether individuals recognize certain objects, people, or places that they have personally experienced. This test might never make it to U.S. courts, but some law enforcement experts are calling for its increased use as an investigatory tool.

In his charge to the Bioethics Commission, President Obama asked specifically that we address the use of neuroscience in the criminal justice system. In addition, the Bioethics Commission received several public comments from key stakeholders recommending consideration of this topic. For example, the MacArthur Foundation Research Network on Law and Neuroscience sought consideration of the admissibility and weight of neuroscientific evidence in the courtroom, and its application to legal questions about lying and memory, prediction of future criminal behavior, and adolescent conviction and sentencing. The application of neuroscience to the legal system and its potential to alter how we understand free will and blameworthiness are topics of both excitement and concern. The Bioethics Commission includes the topic in this report to examine ethical questions about the use of neuroscience within the legal system and how it can inform our fundamental conception of moral responsibility.
SELECTED REPORTS ON NEUROSCIENCE, ETHICS, AND SOCIETY

The President’s Council on Bioethics’ 2003 report *Beyond Therapy: Biotechnology and the Pursuit of Happiness* reflected on ethical and societal implications of using emerging biotechnologies for modifying behavior, slowing age-related decline, and altering emotions, among others.

The National Research Council’s 2008 report *Emerging Cognitive Neuroscience and Related Technologies* explored neuroscience research areas that might have implications for U.S. national security, including advances in measuring individuals’ mental states and intentions, and development of drugs or technologies that can alter human abilities.

The National Research Council’s 2009 report *Opportunities in Neuroscience for Future Army Applications* focused on research and initiatives that might improve the cognitive and behavioral performance of soldiers.

The United Kingdom’s Royal Society released four *Brain Waves* reports in 2011 and 2012 that reviewed neuroscience and neurotechnology developments and their translation into useful applications, the effect of neuroscience advances on education and lifelong learning, potential military and law enforcement applications arising from neuroscience, and neuroscience’s increasing relevance to the law.

The Nuffield Council on Bioethics’ 2013 report *Novel Neurotechnologies: Intervening in the Brain* considered the potential benefits and harms of novel neurotechnologies and outlined an ethical framework to guide their future development, regulation, use, and promotion.

Singapore’s Bioethics Advisory Committee (BAC) is considering current and developing neuroscience research and its ethical, legal, and societal implications. BAC expects to release its final report and recommendations in the near future.

Background and the Promise of Neuroscience

The human nervous system is an exceptionally intricate structure that regulates every organ and tissue of the body. It enables countless automatic functions that we share with other living creatures, including life-sustaining bodily activities, locomotion, and perception of the environment. In addition, it is the locus of more complex human attributes of language, empathy, judgment, and mood. Philosophers, scientists, and other thoughtful human beings have sought to understand the brain and the basis of consciousness, cognition, and volition for centuries.

**NEURONS IN THE BRAIN**

Weighing approximately 3 pounds, the human brain comprises 100 billion neurons and over 60,000 miles of connections that facilitate information transmission. A typical neuron links with several thousand others, resulting in 100 trillion synaptic connections, and allowing for more computing power than any machine. Other cells in the brain provide structure and control the environment necessary for optimal brain function, defend against infection, and support tissue remodeling and repair. Electrical signals are passed between neurons through molecules called neurotransmitters. More than 100 neurotransmitters exist, and most drugs that affect brain function work by altering their levels or function.

The Promise of Neuroscience Research

One of the greatest illustrations of the promise of neuroscience—and a priority for many who study it—is its ability to better the human condition and alleviate suffering. In the original charge to the newly formed National Institutes of Health (NIH) in 1952, Congress defined the purpose of publicly funded biomedical research as follows: “[T]o help provide the practicing physicians of this nation—and of the world—with better means for ameliorating physical suffering and emotional imbalance, for prolonging human life, and for making all the years of that span more useful both to the individual and to society.”

These founding principles are central to neuroscience. Our nervous system determines our capacity to cope with suffering; adapt to adverse conditions; regulate stress, disappointment, and loss; remember past events; plan for the future; and improve the lives of our loved ones. Neurological disorders also are a substantial source of morbidity and mortality. The impact of neurological disorders and psychiatric conditions can be quantified in disability-adjusted life years (DALYs), which account for years lost to disability and premature death (Table 1). They are common and costly and often have devastating impact. More than one billion individuals globally, and millions in the United States, suffer from neurological disorders, with estimated health system and lost productivity costs of over 760 billion dollars a year in the United States alone. Furthermore, as our population ages, the prevalence of age-related neurologic disorders, including neurodegenerative diseases, such as Alzheimer’s disease and Parkinson’s disease, will increase.

Historically, clinicians struggled to diagnose and treat neurological disorders. Before the late 1970s, clinical neuroscience focused on relating a patient’s symptoms with tissue changes observed at autopsy. As diagnosis of neurological problems improved, their future course could be predicted with varying degrees of accuracy, but in most cases, clinicians were unable to intervene in any meaningful way.
Table 1: Global disability-adjusted life years (DALYs) and number of annual deaths for selected neurological disorders and psychiatric conditions in 2010

<table>
<thead>
<tr>
<th>Disorder</th>
<th>DALYs†</th>
<th>Deaths‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low back and neck pain</td>
<td>116,704,000</td>
<td>–</td>
</tr>
<tr>
<td>Cerebrovascular diseases (stroke)</td>
<td>102,232,000</td>
<td>5,874,200</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>17,429,000</td>
<td>177,600</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>1,075,000</td>
<td>18,200</td>
</tr>
<tr>
<td>Alzheimer’s disease and other dementias</td>
<td>11,349,000</td>
<td>485,700</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>1,918,000</td>
<td>111,100</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>14,999,000</td>
<td>19,800</td>
</tr>
<tr>
<td>Depression</td>
<td>74,264,000</td>
<td>–</td>
</tr>
<tr>
<td>Alcohol use disorders</td>
<td>17,644,000</td>
<td>111,100</td>
</tr>
<tr>
<td>Drug use disorders</td>
<td>19,994,000</td>
<td>77,600</td>
</tr>
<tr>
<td><strong>% of total for 10 selected neurological disorders and psychiatric conditions</strong></td>
<td><strong>15.2%</strong>*</td>
<td><strong>13.0%§</strong></td>
</tr>
</tbody>
</table>


*DALYs from 10 selected conditions divided by DALYs from 291 conditions studied; ‡deaths from 10 selected conditions divided by deaths from 235 causes.

Advances in neuroscience over recent decades led to development of treatments for certain debilitating neurological disorders. For example, multiple sclerosis, the most common cause of nontraumatic neurological impairment among young and working-age adults, was previously untreatable, and the vast majority of affected individuals developed irreversible neurological disability. The past two decades, research has yielded more than 10 FDA-approved therapies, dramatically improving the outlook for affected individuals.
Traumatic brain injury (TBI), a disruption of normal brain function caused by impact from an external force, is the world’s leading cause of death and disability among children and young adults.\textsuperscript{39} Ongoing clinical trials are testing different strategies to intervene during the acute phase of TBI.\textsuperscript{40} Harnessing novel neurotechnologies, including electrical stimulation and neuroimaging techniques, might provide new or improved ways to diagnose and treat TBI.\textsuperscript{41} Additionally, advances in neuroimaging might aid in assessing awareness among individuals in a minimally conscious state caused by brain injury.\textsuperscript{42}

An estimated 50 million individuals worldwide have epilepsy, a neurological disorder that is a major cause of morbidity and mortality and can lead to stigmatization.\textsuperscript{43} Contemporary neuroscience led to development of several dozen effective drugs to treat epilepsy. In certain cases, neuroimaging that measures electrical activity of specific brain regions enables clinicians to identify culprit areas of the brain, potentially leading to new treatment options.\textsuperscript{44}

Many neurodegenerative diseases—including Alzheimer’s disease, Parkinson’s disease, amyotrophic lateral sclerosis (Lou Gehrig’s disease), chronic traumatic encephalopathy (resulting from repeated brain trauma), and Huntington’s disease, among others—are associated with accumulation of clumped proteins in degenerating nerve cells.\textsuperscript{45} Evidence indicates that misfolding of culprit proteins into incorrect shapes might be responsible for the degenerative process in each disease.\textsuperscript{46} This misfolding might be genetic or triggered by injury, such as by trauma or infection. Measuring and imaging these abnormal proteins directly and noninvasively in patients is now becoming possible. These advances create a remarkable opportunity to develop effective therapies for neurodegenerative diseases and to monitor their effectiveness in the living brain.

Technological developments in neuroscience offer new hope for individuals with debilitating neurological disorders. For example, deep brain stimulation (DBS) can enable motor function by eliminating disabling tremors, improving the lives of many thousands of individuals with Parkinson’s disease.\textsuperscript{47} In addition, DBS offers promise for other disorders, including psychiatric conditions.\textsuperscript{48}
Deep brain stimulation (DBS) is a surgical intervention that delivers rapidly fluctuating electric current to deep brain structures. The exact mechanism of DBS action is unknown, but it remains a viable treatment option for certain disorders.

DBS is approved by the U.S. Food and Drug Administration for treating Parkinson’s disease, other movement disorders, and treatment-resistant obsessive-compulsive disorder. Research indicates that DBS might be effective in treating neurological disorders and psychiatric conditions, such as depression, epilepsy, and chronic pain. DBS also has potential to alter individuals’ behavior. Because of the invasive nature of the procedure, clinicians and patients consider DBS only after failure of nonsurgical interventions.


In recent years the practice of neuroscience has improved and changed dramatically, fueled by a proliferation of enabling technologies that are transforming our understanding of brain function in health and disease. These technologies include neuroimaging, molecular and cellular neuroscience, medicinal chemistry, and bioinformatics, among others. Sophisticated techniques and technologies have improved prospects in diagnosis, prevention, and treatment of neurological disorders and psychiatric conditions. They have also been central to contributing to a deeper understanding of the human brain, consciousness, behaviors, motivations, intentions, and our sense of self.
**Recent Advances in Neuroscience**

In 2013, scientists developed a method to preserve the integrity of biological tissue while improving its visibility under the microscope. The method replaces dense, irregularly arranged lipid molecules present in normal biological tissue with a hydrogel that allows light and other molecules to pass through easily. The resulting tissue-hydrogel hybrid is transparent, facilitating visualization of large networks of neurons and their connections in the whole, intact brain.

In 2012 and 2013, two teams of scientists demonstrated methods to create a false memory association in mice. Using different techniques and tools, both teams of researchers activated neurons that corresponded to the memory of a foot shock. The scientists were able to reinstate the fear of a foot shock even if the mouse was not placed in the original context in which the foot shock occurred.

In 2011, scientists developed a decoder able to identify and reconstruct moving images viewed by a human participant. Researchers first presented participants with a movie clip and simultaneously measured their brain activity through functional magnetic resonance imaging (fMRI). The decoder, a model of the brain’s visual system, used fMRI data from each individual to produce a primitive reconstruction of the movies seen by the participant.


**The BRAIN Initiative**

The BRAIN Initiative, a White House Grand Challenge, represents substantial federal investment in neuroscience research and a unique opportunity to advance the fruits of neuroscience and its potential to improve human health.

The BRAIN Initiative supports public and private research to develop tools and technologies to understand the human brain. Advances in the ability to map the anatomy and activity of neurons and their connections in different regions in the brain create an opportunity to decipher how groups of neurons and their connections function normally and in disease states. These networks are critical to understanding neurodegenerative disorders. Additional potential exists to understand psychiatric conditions, including depression, schizophrenia, anxiety, post-traumatic stress disorder (PTSD), and addiction, among others. Understanding networks of neurons to decipher the signature
causes of these disorders can help scientists develop more effective diagnostic tools, treatments, prevention, and cures.

Contemporary neuroscience is a rapidly growing, multidisciplinary field, and neuroscientists come from varied academic backgrounds and subfields. Even in such a vast and varied field, we can agree on the value of both the intermediate goal of advancing human knowledge through neuroscience research and the ultimate goal of alleviating human suffering and ameliorating neurological disorders that burden individuals. The BRAIN Initiative, pursued with acumen, persistence, and adequate funding for multidisciplinary teams of researchers, has the potential to help us realize the promise of neuroscience. It supports the development of innovative technologies to improve our understanding of brain function with the ultimate goal to better characterize and treat neurological disorders, including Alzheimer’s and Parkinson’s diseases, depression, PTSD, and TBI. Although more than innovative technology alone is needed to realize the ultimate goal of the BRAIN Initiative, knowledge gained through the use of these new technologies could be key to many diagnostic tools, treatments, preventions, and cures for these and other debilitating neurological disorders.

About this Report

In July 2013, President Obama charged the Bioethics Commission to consider the ethics of neuroscience research and application of neuroscience research findings as part of the BRAIN Initiative. Gray Matters, Vol. 1, released in May 2014, describes many options to achieve integration of ethics with neuroscience research, arguing that a single approach is neither sufficient nor appropriate for all contexts. Integrating ethics and neuroscience is complex: it requires understanding and addressing the unique needs and contexts of individual researchers and institutions. Whatever approach or set of approaches are chosen to integrate neuroscience and ethics, funding and resources are required. For example, scholars have engaged in extensive discussion about one historical model for integration, the NIH’s and U.S. Department of Energy’s Human Genome Project Ethical, Legal, and Social Implications Research Program (HGP ELSI Program). They generally agree that the funding provision for the HGP ELSI Program, which accompanied funds for scientific
research, was a positive aspect of the program. The Bioethics Commission agreed and recommended in *Gray Matters*, Vol. 1 that funding be allocated for ethics integration in neuroscience.\textsuperscript{51}

This report considers in greater detail some of the ethical and societal implications of neuroscience research and its varied applications—implications that integrated ethics and research systems are well-positioned to address. Through deliberation at seven public meetings since receiving the President’s charge, in addition to two relevant meetings undertaken as part of earlier work, and hearing from over 60 experts, the Bioethics Commission identified three controversial topics that exemplify some of the primary ethical considerations at the intersection of neuroscience and ethics. In this report, the Bioethics Commission explores these topics: cognitive enhancement and beyond, capacity and the consent process, and neuroscience and the legal system. The following chapters examine these topics, clarify the current state of the scientific landscape, identify common ground among differing viewpoints, offer an in-depth analysis of ethical considerations, and make concrete recommendations for relevant stakeholders.
CHAPTER 2
Cognitive Enhancement and Beyond
Scientific investigations of the brain and nervous system point the way to new techniques and technologies to modify human neural functioning. Pharmaceuticals, brain stimulation devices, and brain training tools are just some of the neural modification modalities that are in use, under study, or anticipated. Scholars debate the meaning of the term “cognitive enhancement.” Generally, cognitive enhancement describes measures for expanding or augmenting the human capacity to think, feel, react, and remember, potentially “beyond the species-typical level or statistically-normal range of functioning.” Some parties in the debate welcome novel cognitive enhancements as means to human betterment. Others reject the use of biomedical innovations intended to push cognition beyond typical human functioning as threats to moral agency and dignity. What might happen, scholars ask, to traditional understandings of free will, moral responsibility, and virtue if science makes significant advances in the ability to technologically control the mind? In this report, the Bioethics Commission seeks to engage the public in discussion that centers on this important cognitive enhancement debate and moves beyond it to assess a wider array of interventions, technologies, behaviors, and environmental conditions that can affect the functioning of the human brain and nervous system. We use the term “neural modifiers” to refer to this wider array of mechanisms of brain and nervous system change.

Our brains and nervous systems constantly change in response to ordinary daily environmental stimuli, including education, meditation, physical activity, sleep, and diet, among others. Neuroscience helps us better understand the mechanisms of neural change and design novel interventions to alter our brains and nervous systems for a variety of purposes deemed beneficial. Bold new life-changing treatments that improve cognitive, nervous system, and motor function are in use already and more are on the horizon. Researchers seek new approaches to (1) maintain or improve neural health and cognitive function within typical or statistically normal ranges; (2) treat disease, deficiency, injury, impairment, or disorder (referred to as “neurological disorders”) to achieve or restore typical or statistically normal functioning; and (3) expand or augment function above typical or statistically normal ranges. In delineating these neural modification objectives, the Bioethics Commission is mindful that they are not always sharply distinguishable. Drawing clear lines between
maintaining or improving function within normal ranges on one hand and expanding or augmenting on the other, or between treating as therapy on one hand and expanding or augmenting as enhancement on the other, can be difficult in both theory and practice. Moreover, it bears special emphasis that what is “normal” forms a bell-shaped curve that encompasses a range of individual differences, and ethical analysis should not uncritically embrace conventional understandings of normal as ideals of healthy bodies and minds.

Our personal, professional, and societal intentions—whether to maintain or improve, treat, or expand and augment the brain and nervous system—are subject to moral evaluation, as are our ultimate actions and omissions. Altering the brain and nervous system is not inherently ethical or unethical. We can imagine that some neural modifications would violate shared values, whereas others would not. For example, individuals report feeling comfortable with the use of neural modifiers for therapeutic purposes, but are less comfortable about healthy individuals using such products. Additionally, although one might view use of prescription medication as medically necessary to treat post-traumatic stress disorder (PTSD), one also might view the overuse of some of the same drugs, and the resulting memory alteration, as controversial threats to personal identity, among other concerns.

Ethical assessment of neural modification requires consideration of who is choosing the modifier, what is being chosen, what its purposes are, who stands to benefit, and who might be harmed. For example, deep brain stimulation (DBS) is currently used to treat motor disorders, such as Parkinson’s disease, and is under investigation for certain psychiatric conditions. If demonstrated to be effective for problems such as depression, addiction, or overeating associated with obesity, DBS could be ethically acceptable if freely chosen by a fully informed adult. In contrast, the intervention would be ethically problematic if a person is coerced into undergoing DBS by a clinician, researcher, or government authority. A pharmacological intervention, such as modafinil, might be used ethically to treat a disorder of sleep and behavior, such as narcolepsy, but might be considered ethically problematic when used recreationally or to further improve a healthy child’s performance in school. Interventions that are available to improve the health of everyone might pass muster, yet those reserved for economically privileged groups would affront fairness and equity.
In this time of rapidly advancing neuroscience, members of the public must be well-informed, both to make educated practical decisions about personal health and wellbeing, and to participate in collective deliberation and decision making about societal applications of novel techniques and technologies relating to the brain and nervous system. The public should be encouraged and empowered to be ethically thoughtful about where responsibility rests for behaviors and conditions that adversely affect neural health, wellbeing, and the capacity for a good life. Society must be prepared to assess novel neural modifiers as they are developed, to ensure that benefits justify risks, and to assure that relevant ethical concerns are considered, including freedom from unwarranted coercion, fair access, and the preservation of moral capacities that define us as human. The Bioethics Commission seeks in this report to describe the landscape of neural modification—including but not limited to cognitive enhancement—dispel common misconceptions, help educate the broader public, and guide an ethical debate that should be based on reliable empirical and scientific evidence.

**Goals and Purposes of Neural Modification**

As scientific knowledge of the human brain and nervous system expands, so too will our ability to intentionally modify our neural functioning. Interventions that change neural function exist—and more are anticipated—enabled by developments in neuroscience related to memory, learning, intelligence, cognition, and motor and sensory function. As outlined above, goals and purposes of neural modification include: (1) maintaining or improving neural health and cognitive function within the range of typical or statistically normal human functioning, (2) treating neurological disorders, and (3) expanding or augmenting neural function. Against a background of neuroscientific discovery, neural modification with these objectives has emerged as a principal topic for bioethics to consider. Neural modification approaches that go beyond ordinary, low-technology activities into the realm of pharmaceuticals and invasive surgery are of greatest ethical concern due to their heightened potential risk, and possibility for more drastic or irreversible change. We are less concerned about caffeinated coffee, healthy sleep habits, and good nutrition, for example, than we are about use of stimulant drugs and DBS implantation, especially where the intervention serves purposes other than maintaining or improving health or therapeutic treatment of neurological disorders.
Maintaining or Improving Neural Health and Cognitive Function

Maintaining or improving neural health and cognitive function within typical or statistically normal ranges is often a primary goal of modifying the brain and nervous system (Figure 1). Interventions and methods that improve individual and public health and social and environmental conditions can promote and maintain brain and nervous system function.\textsuperscript{60} Many ordinary daily activities can qualify as neural modifiers, and many are presumed ethical because they seem natural and safe, and contribute to a good life in the modern world. For example, good nutrition is essential for optimal cognitive function, and many commentators focus on the role of specific nutrients, such as omega-3 fatty acids, in developing cognitive function.\textsuperscript{61} Although the effect of omega-3 fatty acids on cognitive development is debated, some evidence indicates that supplementing this nutrient among selected populations, such as infants, might be beneficial for several measures of cognitive development.\textsuperscript{62}

\textbf{Figure 1: Maintaining or Improving Neural Health and Cognitive Function within Typical or Statistically Normal Ranges}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Maintaining or Improving Neural Health and Cognitive Function within Typical or Statistically Normal Ranges}
\end{figure}

\textbf{A)} An individual with low (but statistically normal) cognitive function uses a neural modifier to improve cognition.

\textbf{B)} An individual takes a dietary supplement to maintain healthy cognitive function.
Modification of other health behaviors, such as sleep and exercise, also can contribute to maintaining cognitive function. Public health interventions, such as lead abatement programs, can remove impediments to optimal cognitive functioning on the population level as well.

At the individual level, activities such as getting a full night’s sleep, exercising, and maintaining a nutritious diet contribute to brain health and function. Core literacy and numeracy skills and devices, such as computers, calculators, and smartphones, increase our ability to access and process information, concentrate on or attend to information, and communicate with others. Even the caffeine in morning coffee and afternoon tea might qualify as neural modification for the boost to alertness they provide. A cleaner and more pollution- and toxin-free natural environment, along with healthy residential living and working conditions, can contribute to sound neural function. Achieving some of these improved conditions for all comes with a considerable economic price. Yet, as the Bioethics Commission has previously reported, deliberate neglect of the societal background conditions needed for public health is costly, not just in economic terms; neglect of these well-understood improvements in favor of focusing exclusively on new and emerging intervention technologies runs afoul of principles of public beneficence and justice.

With regard to neural health, the relevant background conditions are numerous and therefore provide a range of options worth considering alongside and in addition to investments in neuroscience; these include investments in basic literacy, numeracy, science education, and ethical literacy, as well as environmental protection, and residential and occupational safety.

More controversially, some scholars have suggested that drugs or devices might be developed that can improve cognitive function within statistically normal ranges; however, this is a matter of considerable debate. If and when they become available, these new higher-technology strategies to improve
neural function within the normal range might be more controversial than existing methods. To draw an analogy, recombinant human growth hormone is a safe and readily available drug, approved by the U.S. Food and Drug Administration (FDA) to treat growth hormone deficiency and other growth disorders. Growth hormone also has been used to increase the height of children substantially shorter than the mean for their age and sex, but who are nonetheless “normal” and do not have a diagnosable disorder. The use of this drug has raised concerns about whether it is morally acceptable to use drugs to alter the natural, non-diseased state. Other debates around this issue include whether use of human growth hormone for this purpose should be funded, and by whom.

A similar debate exists in the cognitive enhancement literature. On the one hand, some scholars argue that, if drugs and devices become available that can improve cognitive or other neural function within statistically normal ranges, there is nothing per se unethical about their use. These scholars suggest that novel, high-technology strategies to improve cognitive performance are not inherently different from older, more familiar methods, such as education and calculators. For these scholars, novel interventions do not raise new normative questions although they might raise concerns about safety, distributive justice, or unintended consequences. On the other hand, other scholars contend that drugs and devices used for this purpose reflect a desire to control or master human nature. This desire for control might erode our appreciation for natural human powers and achievements. Neural modification that improves functions relative to a standard distribution is controversial. Some scholars assert that if access to neural modification is universal, gradually increasing average capabilities might fail to improve the relative wellbeing of the least well off. Other scholars view gradual improvement in average neural functioning as a social good worth pursuing. Additional ethical concerns about gradually increasing mean levels of neural functioning include the reduction of diversity (e.g., if neural modification is targeted iteratively to those on the lower end of the normal distribution), and the specter of unintended consequences.

As we reflect on the potential of the BRAIN Initiative and other neuroscience advances to alleviate human suffering, it is worth emphasizing that, if we as a society and individuals have the will, we already have some practical and ethical ways to make substantial improvements in human neural functioning, even
before the neuroscientific advances enable us to do more. Correspondingly, the added knowledge gained by such endeavors as the BRAIN Initiative will need to be supplemented by the social will to invest the resources that are needed to put that knowledge to use in bettering the human condition.

**Treating Neurological Disorders**

A major goal of modifying the brain or nervous system is to treat neurological disorders. Illness, mishap, and aging can result in considerable impairment. Some neural modification is aimed at enabling persons to think, feel, or remember substantially better than they otherwise could.

In the United States and worldwide, the burden of neurological disorders is high and projected to increase considerably in future years. Neurological disorders and their sequelae are estimated to affect as many as a billion individuals globally, including millions in the United States. Furthermore, these disorders occur among all age groups and geographical regions.

For example, Alzheimer’s disease can cause severe memory loss and personality changes. It is the most common neurodegenerative disease in the United States and affects an estimated 5.2 million individuals. Parkinson’s disease impairs motor function and in certain cases can cause dementia. It affects an estimated one million Americans, even striking those who are young. Congenital epilepsies and cerebral palsy can be so severe as to interfere with childhood learning and adult employment. Epilepsy, the fourth most common neurological disorder, affects an estimated 2.2 million U.S. residents. Cerebral palsy, the most common childhood motor disability, affects an estimated one in 323 U.S. children. Autism spectrum diagnoses have increased dramatically during the past several decades, with an estimated one in 68 children being affected. Multiple sclerosis, an autoimmune disease of the nervous system and a common cause of neurologic disability among young adults, also has increased in frequency, especially among women. Automobile crash-related injuries, workplace injuries, combat injuries, and sports injuries leave many individuals without normal brain function, limbs, sight, or hearing. Traumatic brain injuries (TBI), most often caused by car crashes, sports injuries, injuries from military combat, and falls, are the leading cause of death and disability among persons aged 1 to 44 years. Approximately 5.3 million suffer from long-term disability as a result of TBI.
In addition, millions of persons living in the United States are affected by psychiatric conditions, including schizophrenia, bipolar disorder, personality disorder, and major depression. Approximately one in four adults (61.5 million) experiences mental illness in any given year. Approximately 13.6 million individuals live with an ongoing serious mental illness. These disorders can involve disorganized thinking, emotional pain and suffering, extreme emotional dysregulation, and limited self-control, all of which can affect pediatric populations as well as adults. Many serious addictions also are understood today as treatable mental and behavioral health conditions.

The financial cost to society of treating mental and behavioral health conditions is high. For example, the United States spends 113 billion dollars annually on mental health treatment, and serious mental illness is estimated to cost 193 billion dollars in lost earnings per year. But even greater, although in some respects incalculable, are the personal and social costs of caring for those affected by these conditions.

Helping affected individuals to be well again or for the first time is vastly appealing as a public health priority and an ethical imperative of beneficence and justice. As discussed above, one of the primary goals of neuroscience is to relieve human suffering and better the human condition.

Expanding or Augmenting Neural Function

A third major goal of changing the brain and nervous system is to expand or augment cognitive and motor function. The literature refers to many purposes of neural modification to expand or augment function, including cosmetic enhancement, enhancement for the sake of competitive advantage, moral enhancement, and transhumanism (a movement seeking to use technology to radically enhance the human condition). What counts as expanding or augmenting a human trait is contested and subject to debate. Some groups are excited about the possibility that society could be improved by raising the baseline of what is considered within normal human limits. At its most fanciful, this goal is associated with the idea of a future in which human beings will be equipped with functioning more advanced than the typical or statistical norm, such as unyielding stamina or perfect recall by way of bionic men and women, cybernetic organisms, and transhumanism. In a more realistic form, the goal of extending or augmenting neural function is associated with modest cognitive enhancement.
The term “cognitive enhancement” is defined in different ways by the scholars who discuss it, but most uses refer to interventions such as pharmaceuticals, technological devices, and surgeries that improve abilities to think, feel, and remember. Cognitive enhancement is controversial as applied to interventions made available to healthy individuals functioning within normal limits. For example, using psychedelic drugs, including psilocybin and lysergic acid diethylamide (LSD) to achieve what is believed to be spiritual transcendence has been controversial, and criminal law prohibits their use. Ethical concerns about justice and fairness, among others, are raised by prescribing stimulant medications like methylphenidate (Ritalin®) to enable a healthy child without learning disabilities to achieve a higher score on a standardized test. But concerns also are raised by using propranolol to relieve stage fright among performers and to dampen the traumatic memories of survivors of war or sexual assault. These concerns include the ethics of performing better through chemistry, the potential for over-medicalization, and “subsequent exploitation by the pharmaceutical industry.” Further, when we consider altering our memories, we trigger concerns at the core of defining one’s self. Individuals’ memories guide their narrative identities, and modifying them might alter what individuals believe to be true about themselves. Enhancements praised by some observers as improvements are condemned by others as assaults on human dignity, autonomy, and moral virtue.

Intentional modifications of the function of the brain and nervous system, including those debated under the rubric of “cognitive enhancement” raise important ethical questions with implications for clinical care, biomedical research, public policy, and the law. Pharmaceuticals are among the most discussed neural enhancers, but reliable data on the actual prevalence of use of pharmaceutical interventions for cognitive expansion or augmentation are elusive. Few data exist on the prevalence of use of novel pharmaceuticals to boost cognition. The most recent National Survey on Drug Use and Health from the Substance Abuse and Mental Health Services Administration reported that 6.4 percent of college students aged 18–22 reported using combined levo- and dextroamphetamines (Adderall®) non-medically. Survey participants reported using these drugs to improve concentration, focus studying, and increase alertness. An online poll of journal readers conducted by Nature found that one in five of 1,400 respondents had used Ritalin®, Provigil®, or beta-blockers for non-medical purposes (stimulating
focus, concentration, or memory). These results must be interpreted with caution because they were drawn from a small, nonrepresentative sample and are limited by selection bias; however the data demonstrate attempts by healthy individuals to engage in pharmaceutical forms of cognitive enhancement.

Scholars have raised concerns about the available data on the prevalence of cognitive enhancement. Individuals generally are reluctant to admit using prescription drugs for non-medical reasons, making it difficult to estimate prevalence. Published surveys focus on specific populations (e.g., college students) rather than the general population. This limited focus can undermine claims about the widespread use of stimulants for cognitive enhancement. Surveys often do not capture accurately whether the drugs are being used for cognitive enhancement or other purposes. For example, many surveys ask respondents about non-medical use, but do not distinguish between non-medical use for enhancement and recreational purposes. Inadequate data on the prevalence of use of novel forms of cognitive enhancement can lead to under- or overestimates. Generalizing findings from selected groups more likely to use cognitive enhancing drugs can lead to overestimation of prevalence and a misleading sense of urgency regarding the need for regulatory frameworks and policies. Better evidence of prevalence is essential for a careful and considered assessment of the ethical challenges raised by novel forms of cognitive enhancement. Clearly, as neuroscience advances, individuals will continue to experiment with new ways to enhance their cognition, and we will learn more about the benefits and risks of different modalities of enhancement.

Neuroscience research demonstrates that some drugs and brain stimulation devices can have modest enhancing effects on some cognitive abilities in healthy individuals under certain conditions. However, the size of these effects and their generalizability to real-world (rather than laboratory) settings remain uncertain. These include the stimulant drugs Ritalin® and Adderall®, used in the treatment of attention deficit hyperactivity disorder (ADHD); Provigil®, prescribed for disorders of excessive sleep; and such methods as transcranial magnetic stimulation and transcranial direct current stimulation (tDCS). For example, one review of Provigil® and Ritalin® use for cognitive enhancement states that “expectations regarding the effectiveness of these drugs exceed their actual effects.”
This observation illustrates one problematic consequence of hype: Although the actual enhancing effects of these interventions might be limited and can vary among individuals, widespread discussion and debate around them can contribute to unfounded and inflated expectations. In addition, a lack of discussion about unrealistic expectations of benefits and actual risks of side effects can result in more harm to individuals. Novel and still largely experimental modalities of cognitive enhancement raise ethical considerations not relevant when considering more conventional modes because we know and understand less about their potential benefits, if any; their potential risks, including long-term effects; and the potential societal impact of their widespread use likewise are underexplored.

Individuals use many of these interventions therapeutically, and they have well-documented safety profiles in the medical context. However, in the context of healthy individuals, a lack of reliable long-term data on the use of pharmaceuticals for nontherapeutic purposes means that the potential risks are unknown. Some potential risks, such as the risk of dependence, are of particular concern. The risk-to-benefit profile for healthy individuals using pharmaceutical cognitive enhancement is very different than it is for individuals using these interventions for therapeutic purposes, especially given the lack of evidence on the cognitive enhancing effects.

Even less is known about the risks and long-term effects of novel cognitive enhancers among children and adolescents. Unknown long-term risk is of concern, especially in the context of children’s developing brains. Both the use of prescription drugs to treat ADHD among children and the use of prescription drugs for cognitive enhancement among children are increasing. Four percent of 12th graders in the United States reported use of Adderall® and 2 percent use of Ritalin® when asked what amphetamines they have taken during the previous year without a doctor’s orders. Pediatric use without a clinician’s supervision is concerning because ADHD medications can cause
rare but serious cardiovascular side effects, especially in children with a family history or known risk factors.\textsuperscript{110} Potential also exists for abuse or dependency.\textsuperscript{111} Research with animals demonstrates that Ritalin\textsuperscript{®} can induce changes in neuronal morphology, and can cause behavioral changes in adolescent mice and rats, that persist through adulthood.\textsuperscript{112} A position paper endorsed by the American Academy of Neurology, Child Neurology Society, and American Neurological Association recommends that physicians refrain from prescribing drugs for nontherapeutic purposes among children.\textsuperscript{113} More research is needed about the neurodevelopmental and behavioral implications of stimulant drug use in children and adolescents.\textsuperscript{114}

Individuals also can purchase some neural modifiers directly from companies without a medical intermediary, for example, over-the-counter dietary supplements or tDCS devices. However, other modifiers (e.g., prescription medications) require a treating clinician. Although consumers generally are responsible for learning about and taking responsibility for any risks involved in using an intervention purchased without a health care professional’s oversight, when clinicians are involved, they have a responsibility to consider patients’ best interests and warn them about potential risks of taking prescription medication—especially for non-medical use.

No comprehensive or agreed-upon guidance or best-practices document is available to advise stakeholders (including clinicians, employers, parents, educators, and patients, among others) about potential benefits and risks inherent in using a neural modification intervention. Educators, parents, clinicians, and employers in highly competitive fields would benefit from having expert, evidence-based advice when faced with student or employee use of neural modification interventions.
Novel means of enhancing cognition that rely on advances in neuroscience, although still in their infancy, are likely to become more effective in the future. Neuroscience research has much to contribute to our understanding of whether these interventions are effective at enhancing cognition and the level of risk they impose for healthy individuals, including children. Thus, now is an opportune time to engage the public and all relevant stakeholders in a discussion about the ethics of novel neural modifiers.

Ethical Analysis

Modifying the brain and nervous system is not inherently ethically problematic. Individuals use a wide range of substances, processes, and interventions to modify the brain and nervous system, including high-quality nutrition, meditation, education, drugs, and devices. Society must evaluate the ethical concerns of specific means of neural modification individually, including those labeled cognitive enhancements, to determine whether and why they are potentially problematic. Scholars characterize ethical issues raised by cognitive enhancement into multiple clusters, including “freedom [and] autonomy, health [and] safety, fairness [and] equity, societal disruption, and human dignity.” The debates about cognitive enhancement include many of the ethical concerns raised by neural modification more generally, including the importance of facilitating healthy development and wellbeing; respecting moral agency; informed consent to medical procedures and research; minimization of risk; public education and deliberation; equity and access across all demographic groups; and the reduction of disadvantage, suffering, and stigma associated with neurological disorders.

Benefiting Individuals and Society

Principles of beneficence and non-maleficence require taking steps to promote the health and wellbeing of oneself and others and avoiding harm. With these
ideals in mind, society has a responsibility to consider the potential benefits and risks of neural modifiers. Safe and effective neural modification that can alleviate suffering and the felt burdens of human impairment are generally considered morally acceptable. In fact, development and use of neural modifiers to maintain or improve health or treat disease represents one of the primary goals of neuroscience research and advances both individual and public beneficence.

Neural modifiers that maintain or improve health, or treat neurological disorders might not be universally applicable to all members of society, but can be beneficial to individuals who need an intervention. For example, correcting a vitamin B12 deficiency in an individual might be effective in maintaining healthy cognitive function and staving off cognitive decline. Or, in a clinical context, a psychiatrist might prescribe the drug clonazepam (Klonopin®) for a patient suddenly experiencing severe panic attacks after a home invasion. If the drug reduces the frequency and severity of panic symptoms, the patient might return to the quality of work and home life enjoyed before symptom onset. Nothing is inherently unethical about voluntarily taking a dietary supplement known to be safe or a prescription medication shown to be safe and effective by adequate scientific research to maintain or improve health or treat a disease or disorder.

Neural modifiers that improve cognitive ability also offer considerable instrumental benefits. Cognitive abilities can influence important outcomes for individual lives, including success at work, earning potential, likelihood of experiencing social and economic difficulties, and overall health. On a societal level, widespread improvements in cognitive function might produce collective benefits, such as economic gains or improved safety from error reductions in high-risk professions and the military. Certain neural modifiers, such as education and those that treat or prevent disease, could be morally required in societies capable of delivering them because of their potential to advance both individual and public beneficence. For example, in the United States, where access to elementary and secondary education is a legal mandate, enhancing cognitive capacities through schooling is a priority.

Neural modifiers can thus provide substantial individual and societal benefits through interventions that nurture healthy brain development, counteract the effects of impediments to development, and promote optimal functioning.
Justice and Fairness

Concerns about justice and fairness related to neural modifiers that enhance cognition and other functions of the brain and nervous system arise in two distinct ways. First, an individual with more—or enhanced—cognitive abilities, for example, might have an advantage relative to others; in this sense, cognitive ability is a positional good, in that it confers an advantage on some individuals only if others do not have the same good. Cognitive enhancement raises the concern that those who have access will gain an unfair competitive advantage over those who do not.\textsuperscript{124} If safe and effective novel forms of neural modification are available only to those who are already advantaged (by wealth or social capital), limited availability might exacerbate existing inequalities.

Justice and fairness requires not only equitable distribution of the benefits of neural modifiers, but it also requires attention to the distribution of their burdens and risks. For example, early study and use of neural modifiers might find that they are effective in the short term, but cause negative consequences in the long term. The burdens and risks of understudied neural modifiers must not fall unfairly on certain groups or individuals.

Second, neural modifiers thought to alter cognitive and other neural functions can offer nonpositional benefits (i.e., benefits that are inherently valuable, not because they provide a competitive advantage over others). For example, having a higher earning capacity is beneficial to individuals not (or not only) because it gives them an advantage over others, but because it provides them with means to secure better living conditions and a greater range of opportunities. Similarly, knowledge is considered a good in itself, and access to more knowledge can be inherently valuable. Here, the concern about justice is not whether access to the means for elevated brain and nervous system function confers unfair advantages, but rather whether the distribution of safe and effective neural modifiers can promote justice by providing individuals with a greater range of opportunities and enabling them to participate more fully in society.\textsuperscript{125}

The nonpositional individual and societal benefits of neural modification support pursuing modifications collectively, rather than limiting access to a privileged few. Neuroscience research on the effects of novel neural modifiers can contribute to our understanding of how these interventions can be
distributed justly. For example, evidence demonstrates that the effects of certain pharmaceutical cognitive enhancements depend on baseline cognitive functioning. Individuals with lower levels of baseline functioning appear to experience a greater improvement than those at a higher baseline level.\(^{126}\) If these results are borne out by further research, cognitive enhancement interventions could be used to reduce inequities between the cognitively advantaged and disadvantaged, for example, by reducing gaps in educational achievement. Some scholars argue that if cognitive enhancement and other neural modifiers could reduce existing inequities, then justice requires interventions.\(^{127}\) This might prove to be so. At the very least, new forms of safe and effective interventions that deliver real advantages to those who use them should not be distributed so as to exacerbate or amplify existing inequities.

Access to low-technology neural modifiers, such as educational enrichment, test preparation courses, or adequate childhood nutrition, raise justice and equity concerns parallel to those raised by novel neural modifiers. Societal tolerance of inequity in access to other crucial goods does not make inequity right, nor should it hamper our efforts to reduce or eliminate inequity where we can.\(^{128}\) If safe and effective novel forms of cognitive enhancement become available, they will present an opportunity to insist on a distribution that is fair and just. While not eliminating all other less tractable forms of injustice in the distribution of neural health and wellbeing, it is possible to ensure that any new forms of safe and beneficial neural modification do not worsen those injustices.

**Moral Agency and Human Dignity**

Moral agents are individuals capable of acting freely and making judgments for which they can be praised, blamed, or held responsible. Respect for human dignity has grounded longstanding ethical prohibitions against coerced uses of drugs and devices to alter the brain and nervous system.\(^{129}\) In addition, some scholars contend that cognitive enhancement and other neural modifications also pose a potential threat to moral agency and human dignity.\(^{130}\) Enabling individuals without specific impairments to achieve higher levels of cognitive function is ethically controversial. Scholars question whether humans should exercise so much control over the natural world, and debate where to draw the line.\(^{131}\) On this view, advances that vastly improve human beings cross ethical lines by risking the creation of not better humans but transhumans.
Use of pharmaceuticals to improve alertness, attention, mood, and happiness also raise concerns about morally legitimate paths to success and wellbeing. Some scholars consider achieving success with the help of a pill akin to cheating or taking the easy way out, because they believe success is supposed to be the result of personal effort and hard work. According to this objection, some forms of neural modification offer only false visions of human achievement. This type of success might be valuable for the immediate outcome, but cannot be considered the kind of achievement that results from personal will and exertion. From this perspective, success is as much about how goals are achieved as achieving them.\textsuperscript{132}

Similarly, some scholars contend that happiness and wellbeing are supposed to be rewards of virtue and good character, not an outcome of medication.\textsuperscript{133} Although it can be deeply upsetting and profoundly life-changing to live with traumatic memories, some view medications to dampen memories as problematic because they could prevent individuals from coming to terms with their lives as continuous subjects of both good and bad experiences.\textsuperscript{134} From this perspective, neural modification, particularly through pharmacological management, threatens to provide only “fraudulent happiness.”\textsuperscript{135} Yet, from another perspective, ethical merit might exist in “fraudulent happiness” that enables individuals to be functional parents, providers, and engaged citizens.

In contrast, others view the practice of novel neurotechnologies being used to enhance humans as technological progress and innovation. Scholars evaluate some forms of neural modification for their potential to be used for moral enhancement. Drugs that free us of rage, impulsivity, and aggression might enable us to participate successfully in the moral community. For example, some research results indicate that oxytocin can promote generosity, and other pharmacological substances have been reported to increase cooperation.\textsuperscript{136}

Importantly, the empirical evidence supporting the possibility of moral enhancement is thin, and interpreting results in terms of moral enhancement has been criticized by both those internal and external to the scientific community.\textsuperscript{137} Some scholars question whether we would be morally better at all if only through use of a drug—our conduct would be the result of a will controlled by the external and artificial stimulus of a pharmaceutical rather than will disciplined through effort. In contrast, others point out
that, although technological moral enhancement is only a distant prospect, it can serve as a complement to, not a replacement of, traditional social and educational modes of moral improvement.\(^\text{138}\)

**Importance of Public Education and Deliberation**

In its 2010 report on synthetic biology, *New Directions: The Ethics of Synthetic Biology and Emerging Technologies (New Directions)*, the Bioethics Commission recognized the importance of informed and reasonable public debate about potentially controversial issues in science, technology, and ethics, emphasizing the importance of an informed public to facilitate democratic deliberation.\(^\text{139}\)

In this report, it recognizes that the debate about cognitive enhancement and other novel neural modifiers can be fraught with exaggeration and misinformation. Scientists, the media, policymakers, and other stakeholders often hype or inaccurately portray facts.\(^\text{140}\) A deeper and more accurate understanding of relevant evidence, realistic potential, and the true and pressing ethical concerns surrounding controversial topics is essential.

In *New Directions*, the Bioethics Commission recommended that a mechanism be created to fact check the diverse claims made about advances in synthetic biology.\(^\text{141}\) Similarly, in the cognitive enhancement and novel neural modification arena, a mechanism analogous to FactCheck.org would be a useful tool to facilitate a more informed consumer and public.\(^\text{142}\)

Educated public debate about science helps to air all relevant perspectives, and participation by the scientific community helps maintain an educated public.\(^\text{143}\) This fosters democratic deliberation—collaborative decision making that embraces respectful debate of opposing views and active participation by members of the public. The Bioethics Commission urges our society to uphold this particularly important ethical principle to maximize the potential that emerging neuroscientific discoveries will be well-understood and used for the betterment of the human condition.\(^\text{144}\)

**Recommendations**

Through debates about cognitive enhancement, neural modification has captured the public’s imagination. Neural modification includes basic strategies already demonstrated to improve brain and nervous system function. These evidence-based strategies include healthy diet, exercise, sleep, and education.
Novel neural modification methods foster hope that we soon will be able to ameliorate the symptoms of depression, curtail addiction, and prevent dementia-induced cognitive decline, among other health benefits. In addition, contemplating novel methods to improve such functions as learning and memory in school or performance in competitive professions is truly exciting.

Public discourse around neural modification, including cognitive enhancement, reflects fascination, but also raises ethical concerns. Although unjust access to beneficial interventions could exacerbate social inequities, some interventions also could promote equity by closing existing gaps. Fully appreciating any risks in addition to potential benefits inherent in using neural modifiers is imperative. Guidelines for practitioners and relevant stakeholders can help guide use of brain and nervous system interventions and their potential risks and benefits in diverse circumstances.

Several well-known lifestyle interventions, such as adequate sleep, exercise, and nutrition, are associated with improved neural function. Similarly, public health interventions, such as lead paint abatement, can help prevent the negative impact of environmental exposures on neural development and function. These behaviors and conditions can maintain and improve neural health and might be safer and more effective than those offered by novel neural modifiers, some of which might have minimal benefits and uncertain risks.

**Recommendation 1: Prioritize Existing Strategies to Maintain and Improve Neural Health**

In addition to developing new drugs and devices to maintain and improve neural health, funders should prioritize and support research on existing, low-technology strategies, such as healthy diet, adequate exercise and sleep, lead paint abatement, high-quality educational opportunities, and toxin-free workplaces and housing.

Implementation of evidence-based strategies to maintain or improve cognitive, motor, and nervous system function through behavior or environmental modification should be prioritized over strategies based on novel and often very expensive interventions for which evidence of effectiveness remains uncertain. Funders should allocate resources to study and implement
strategies—both low- and high-technology—that support neural health. For example, a long history of studies demonstrates that education and adequate sleep can improve cognitive function. Neuroscience can help us better understand the mechanisms for learning and sleep, the effects that these activities have on the brain, and how to make them more efficient and effective. A thorough, neuroscientific understanding of known interventions can contribute to implementation of effective strategies for improving neural functioning across society.

Existing treatments for neurological disorders are valuable and can be improved. In addition, emerging neural modification interventions will help reduce the individual and societal burden of neurological disorders. For example, brain-computer interfaces might allow those paralyzed from spinal cord injury, brainstem stroke, Lou Gehrig’s disease, or other disorders to perform complex and flexible movements with a neurally controlled robotic arm. Early research indicates that brain stimulation techniques might alleviate symptoms in wide-ranging neurological disorders and psychiatric conditions, and some FDA-approved brain stimulation devices are currently being used to treat conditions like bipolar disorder, but ethical concerns remain. Safe and effective treatments can improve the lives of millions of individuals living with such conditions.

Recommendation 2: Prioritize Treatment of Neurological Disorders

Funders should prioritize research to treat neurological disorders to improve health and alleviate suffering. This research should consider individual, familial, and public health burdens as well as potential risks, benefits, and long-term effects of specific interventions.

Research on interventions to treat neurological disorders should be prioritized. Funders should allocate resources to study and implement productive, low-technology strategies in addition to developing new treatment interventions.
Although the Bioethics Commission recognizes the need to prioritize the study of both traditional and novel interventions for the prevention and treatment of neurologic disorders, it nonetheless also supports research to better characterize and understand novel neural modification techniques to augment or enhance function. Limited, inconclusive evidence exists for the benefits and risks of stimulant drugs, such as Ritalin® and Adderall®, and brain stimulation methods, such as tDCS, as neural enhancers. In addition, few data are available on the prevalence of the use of neural modification interventions for cognitive enhancement purposes. Most prevalence data for stimulant drug use are limited to specific populations, such as college students, and prevalence surveys generally do not capture why individuals use the drugs non-medically (i.e., whether for cognitive enhancement or other reasons). Only limited anecdotal evidence exists on the use of brain stimulation devices for enhancement purposes.

**Recommendation 3: Study Novel Neural Modifiers to Augment or Enhance Neural Function**

Funders should support research on the prevalence, benefits, and risks of novel neural modifiers to guide the ethical use of interventions to augment or enhance neural function.

The lack or misinterpretation of evidence on prevalence, benefits, and risks can contribute to exaggerated expectations and pronouncements surrounding neural modifiers, making them seem more widespread, effective, or threatening than they are. Ethical analyses must account for limitations of available evidence. Targeted research will provide better evidence to ensure that an accurate message is portrayed to the public about the potential impact of these interventions. This research should consider the prevalence of use in a variety of educational and professional settings, potential risks involved in the use of specific interventions, long-term effects, and effectiveness in real-world settings.

Better evidence is needed on which to base ethical deliberations. For example, concerns about human dignity, distributive justice, and fair access to enhancing interventions will be most salient if and when neural enhancement interventions are demonstrated to be both beneficial and safe in enhancing neural function. Use of neural modification interventions should be supported
by robust ethical deliberation guided by ample evidence on the benefits and risks of each intervention.

One prominent ethical consideration is access to the benefits of neural modification. Limiting access to effective enhancement interventions to those who already enjoy greater access to other social goods would be unjust. It also might deprive society of other benefits of more widespread enhancement—including societal benefits, such as improved civic engagement or greater productivity—that increase as more individuals have access to the intervention.¹⁵⁴ In addition, more widespread enhancement might help to close some gaps in opportunity that are related to neural function, such as educational attainment or employment. These potential benefits support the claim that access to safe and effective enhancement interventions should not be limited to those with financial or other means.

**Recommendation 4: Ensure Equitable Access to Novel Neural Modifiers to Augment or Enhance Neural Function**

Policymakers and other stakeholders should ensure that access to beneficial, safe, effective, and morally acceptable novel neural modifiers to augment or enhance neural function is equitable so as not to compound or exacerbate social and economic inequities.

One way to address concerns about justice is to ensure that proven enhancement interventions are available to everyone or to no one; however, both of these extremes can be unjust. Making enhancements available to everyone, although fair, might simply preserve existing inequities. Making enhancements available to no one, although similarly fair, might deprive individuals and society of the potential benefits the intervention could bring. Although limited, evidence indicates that some cognitive enhancement technologies might confer the greatest benefit on those most in need. If this is demonstrated to be the case, these technologies might reduce gaps in cognitive performance that can have substantial implications for an individual’s social and economic position.¹⁵⁵

Evidence about who benefits most from neural enhancements can guide policies that are sensitive to the contours of social and economic disparities.¹⁵⁶
Enhancement interventions that are demonstrated to be safe and effective should be assessed to determine their potential to affect social and economic disparities. Just as additional evidence is needed to understand what interventions—if any—are beneficial, safe, and effective, additional debate is needed to inform whether specific interventions are morally acceptable.

Clinicians often receive requests to prescribe medications for cognitive enhancement. Some guidance suggests that prescribing medications to adults for the purposes of cognitive enhancement can be considered ethically permissible, yet individual clinicians must decide whether to prescribe the medication to particular patients. These decisions are more ethically complex with regard to children, because children lack legal and ethical consent capacity and are vulnerable to coercion.

Clinicians considering prescribing interventions for neural modification, including cognitive enhancement, should have access to detailed professional guidelines that can help them manage patient requests ethically, especially with regard to children and adolescents. Other stakeholders also would benefit from education and guidance on neural modification interventions. These stakeholders include employers, parents, educators, and professional organizations in fields such as aviation, medicine, and the military, among others, that are associated with on-the-job use of brain and nervous system enhancement interventions.

**Recommendation 5: Create Guidance About the Use of Neural Modifiers**

Professional organizations and other expert groups should develop guidance for clinicians, employers, parents, educators, and patients about the use of neural modifiers and their potential risks and benefits. Medical professional organizations should develop guidelines to assist clinicians in responding to requests for prescriptions for interventions to expand or augment neural function. Clinicians should not prescribe medications that have uncertain or unproven benefits and risks to augment neural function in children and adolescents who do not have neurological disorders.
Widely diverse groups and professional organizations can contribute to the development of guidelines on the prescription and use of neural modification interventions. Detailed guidance and educational materials for stakeholders can guide decisions about how to reconcile the potential benefits and risks of using a particular neural modifier under various circumstances—for example, an individual with age-associated memory impairment or a parent seeking medication for a healthy child to improve unimpaired cognitive function.161

Health care providers are the gatekeepers of many medications requested for enhancement purposes and, through their professional organizations, are well-placed to develop comprehensive guidance on appropriate prescribing practices.162 Professional organizations in other relevant fields, such as education, guidance counseling, aviation, and medicine, should work to develop policies on the use of enhancement interventions. These policies can guide health care providers’ decision making and help to protect individuals from pressure to use enhancement interventions. Policies should inform stakeholders about the ethical concerns that arise with the prescription and use of neural modifiers, including justice, risk, coercion, and respect for human dignity. Guidelines can clarify misunderstandings and prevent ethical missteps.

Generally, clinicians should not prescribe medications that have uncertain or unproven benefits and risks to augment neural function among children and adolescents who do not have neurological disorders.163 The American Academy of Neurology endorses the position that cognitive enhancement with prescription drugs is not ethically justifiable among a healthy pediatric population.164 The pediatric population presents unique ethical concerns: Children lack legal and ethical consent capacity, and clinicians and parents must make decisions that account for children’s developing autonomy, their right to an open future, and their vulnerability to coercion and undue pressure.165 The current state of evidence on novel neuroscience interventions to augment or enhance indicates that, generally, they should not be used among the pediatric population.

* * *

By broadening the discussion of cognitive enhancement to include all forms of neural modification, the Bioethics Commission is expanding the scope of the current debate. Neural modification—to maintain or improve brain health
within typical or statistically normal ranges, treat neurological disorders, and expand or augment neural function—raises a set of ethical considerations, including justice concerns, questions about how to reconcile risks and benefits on a case-by-case basis, and concerns about personhood and moral agency. The Bioethics Commission asserts that cognitive enhancement using novel neurotechnologies exists on a spectrum with other neural modifiers. Thus, it recommends that stakeholders focus on a set of priorities for developing and using neural modifiers, including prioritizing safe and well-studied methods for neural enhancement, prioritizing the development of neural modifiers that have the potential to treat disorders, and conducting more research on the prevalence, benefits, and risks of novel neurotechnologies.
CHAPTER 3
Capacity and the Consent Process
Scientific progress to improve human health requires human participation in clinical research. Contemporary neuroscience research promises to provide important insights into the nature of disease, as well as possible prevention strategies, diagnostic tools, and treatments for a range of increasingly prevalent and often devastating neurological disorders and psychiatric conditions. Much-needed neuroscience research is ongoing and will continue to be conducted on disorders that are often associated with impaired consent capacity. Such research often concerns the very organ responsible for decisions about whether to participate in such research in the first place.

Informed consent, based on the principle of respect for persons, is a foundational tenet of clinical and research ethics.\textsuperscript{166} It is a widely accepted ethical, legal, and regulatory requirement for most clinical research and health care interactions.\textsuperscript{167} To give informed consent, one must have the capacity to provide consent, known as consent capacity. The underlying abilities that consent capacity comprises are debated, but often are thought to include an ability to understand disclosed information, appreciate its significance, and use the information to reason and make and express a choice.\textsuperscript{168} Neuroscientists who conduct research involving human participants commonly work with populations or individuals whose consent capacity might be absent, impaired, fluctuating, or in question.\textsuperscript{169} Similarly, clinicians in many settings encounter patients with impairments in consent capacity. Widely diverse disorders and injuries can affect an individual’s capacity to understand information, consider the benefits and risks of research participation, or reach an informed decision regarding study participation. Neurological disorders, such as head trauma, stroke, dementia, neurological cancers, and metabolic disorders, affect neurological function and can lead to impaired decision-making capacity. Individuals with psychiatric conditions, including schizophrenia or major depression, and those who use psychoactive medications or addictive substances also might have impaired consent capacity. Of note, not all individuals with these conditions have diminished consent capacity or consistently diminished capacity—some affected individuals are capable of understanding information and providing informed consent some or all of the time. In addition, certain novel neuroscientific research interventions can alter participants’ consent capacity. For example,
procedures such as deep brain stimulation (DBS) or electroconvulsive therapy might pose risks to cognitive function.\textsuperscript{170}

Neuroscience research is a principal means of promoting progress and benefiting populations affected by neurological disorders and psychiatric conditions, including those associated with impaired consent capacity. Substantial progress in understanding many of these disorders has been possible because of the advances in research, yet much more needs to be done. This research is ethically challenging because it requires participation of individuals with disorders associated with impaired consent capacity, and voluntary, informed consent is an important ethical tenet in the protection of research participants. To reconcile these competing commitments, such research should only proceed with additional ethical safeguards and protections in place. Protections might include robust initial and ongoing assessment of consent capacity; methods to improve informed consent to accommodate participants’ needs, including audiovisual means and paced verbal instructions; methods to respect assent and dissent when consent capacity is partial or in question; independent consent monitors; limits on risk; clear parameters and procedures for obtaining the permission of a legally authorized representative (LAR) when a participant lacks consent capacity; research advance directives; and stakeholder engagement. In addition, use of research advance directives (a set of written instructions articulated by an individual to appoint a proxy and to direct their involvement in future research) and attempts to mitigate stigma associated with conditions that lead to impaired consent capacity can provide protection.\textsuperscript{171} Clear practices that are well-articulated publicly are needed to protect those with impaired consent capacity while promoting vital neuroscience research.

Grappling with the complex challenges surrounding informed consent and consent capacity requires diverse expertise, including perspectives from

\begin{quote}
“We don’t want to stereotype people based on a diagnosis…[or] disrespect the autonomy of people who are still able to make their own choices. On the other hand, you don’t want to fail to protect vulnerable people who can’t understand the decisions before them. That would be taking advantage of them for the greater good.”

\end{quote}
neurology, psychiatry, psychology, social work, patient advocacy, and bioethics. In addition, the concept of consent capacity, the causes of its impairment and its potential to be restored through therapeutic intervention are areas that stand to benefit greatly from the fruits of neuroscience research. Neuroscience research could help refine our understanding and assessment of decision-making capacity, including consent capacity, and its underlying neurological correlates. Neuroscience research also has the potential to help us understand what abilities and decision making skills are needed for effective consent with individuals affected by a wide range of conditions, and to inform improvements in managing impaired consent capacity. Thus, an ethical analysis of consent capacity is well-suited for this report on neuroscience and ethics.

**Ethical Analysis**

Contemporary neuroscience research presents an opportunity to achieve a deeper understanding of brain-related disorders that represent a major public health burden and have a severe impact on caregivers and loved ones (see Chapter 1: Background and the Promise of Neuroscience Research, above). Some of these disorders, however, are associated with impaired, fluctuating, or diminishing decision-making abilities, which can affect the individual’s capacity to consent. It is vital, wherever possible, to find ways to ethically and responsibly include individuals with impaired consent capacity in neuroscience research, as well as to use neuroscience to better understand the capacities that enable and impede informed consent.

As with many of today’s research protections, concern for participants with impaired consent capacity stemmed initially from revelations of past abuse and mistreatment, such as experiments with institutionalized individuals, and widespread public concern regarding psychosurgery. Against this backdrop, national advisory bodies, researchers, and institutional review boards (IRBs)—committees that review human subjects research—have struggled to both protect against future abuses and not unjustly or unnecessarily exclude potential participants because of their condition or impaired consent capacity or for fear of legal liability. This dual mission—protection and inclusion to ensure that the benefits of research are distributed equitably—shapes many core ethical considerations surrounding capacity, the consent process, and participation in research.
Ensuring Access to the Benefits of Research through Inclusion

Several foundational principles of bioethics—respect for persons, beneficence, and justice and fairness—support the inclusion of participants with impaired consent capacity in neuroscience research, with appropriate protections in place.

Inclusion in research of persons who might have impaired consent capacity reflects respect for persons, which encompasses respect for other forms of agency—in other words, even though individuals might lack autonomy, they might have the ability to express certain preferences or participate in some way in decision-making processes. The principle of respect for autonomy calls on us to respect the abilities of others to reason, come to considered judgments, and make decisions regarding what is best for them.\textsuperscript{174} The broader principle of respect for persons encompasses more than just respect for autonomy, recognizing that all people, including those who are not autonomous, deserve respect. Respectful research practices can take many forms, including thoughtful implementation of additional protections for participants who lack consent capacity.\textsuperscript{175} Providing potential research participants with information about the purpose, prospect of benefit, and risks related to a research protocol enables individuals to take part in deciding whether to participate. Respect for persons is therefore at the core of concerns about consent capacity in research. The principle requires neuroscience researchers to support autonomy and all forms of agency whenever possible.\textsuperscript{176} This includes respect for expressions of agency that, when encountered, reflect meaningful participant values or preferences.\textsuperscript{177} Respectful practices also include facilitating measures, such as research advance directives, which help individuals express their wishes for the future.\textsuperscript{178} In addition, respecting participants who might have impaired, fluctuating, or diminishing consent capacity means making every effort to avoid two equally troubling mistakes: misidentifying capable individuals as incapable, and misidentifying incapable individuals as having consent capacity.\textsuperscript{179}

Research benefits can accrue to affected populations if research participation practices are inclusive. Beneficence calls for efforts to secure the wellbeing of others. Public beneficence confers on society an obligation to advance scientific and technological discovery that can improve public wellbeing.\textsuperscript{180} Neuroscience research can potentially improve the prevention, diagnosis, and treatment of
disorders that can lead to cognitive impairments; however, failing to support neuroscience research on certain disorders because potential participants might have impaired consent capacity can do a disservice to current and future patients.\textsuperscript{181} For example, a 2008 study used surrogate consent to enroll participants with impaired consent capacity to study an investigational treatment that later became the only U.S. Food and Drug Administration (FDA)-approved treatment for acute ischemic stroke. Researchers estimated that, had they not been able to include these participants, the research would have taken four times as long and its validity and impact would have been undermined.\textsuperscript{182}

When research participation practices are inclusive, fair distribution of research benefits is made possible. Justice and fairness requires that the benefits of neuroscience research be distributed equitably across society.\textsuperscript{183} In the previous stroke research example, failing to include persons with impaired consent capacity might have been a disservice to many patients who suffer strokes, because research conclusions would not have been readily generalizable to groups outside the study population.\textsuperscript{184} To address the very conditions that threaten to impair consent capacity, we should do our best as a society to find ethical means wherever possible to include individuals with impaired consent capacity—with appropriate protections in place—in potentially path-breaking research.

\textit{Protecting All Research Participants}

As the history of research ethics reveals, researchers have sometimes selected those with impaired consent capacity for research participation in part because of their greater convenience as a participant pool.\textsuperscript{185} Ethical human subjects research can involve treating people as means as well as ends-in-themselves, whereas, when participants are exploited only to further the interests of others, they are being used as \textit{mere} means.\textsuperscript{186} That is ethically unacceptable. Prohibiting the exploitative treatment of others as \textit{mere} means to scientific advancement is critically important and at the heart of current research oversight. Development of additional ethical protections is meant to help avoid future transgressions.\textsuperscript{187}

Subpart A of the U.S. Department of Health and Human Services’ (HHS) regulations, Protection of Human Subjects (codified at 45 C.F.R. Part 46), also known as the Common Rule, has been adopted by 18 federal departments and agencies that conduct or fund human subjects research and provides standards
for the ethical conduct of federally supported human subjects research. It identifies “mentally disabled persons” as a population that is potentially more vulnerable to coercion or undue influence.\textsuperscript{188} Research ethics often invokes the concept of vulnerability to highlight the unique needs of certain populations who participate in research. Importantly, this framework of vulnerability highlights the ethical goal of providing special protections to those participants who might be more susceptible to exploitation or harm than others as a result of research participation. The concept of vulnerability in research ethics is longstanding and complicated.\textsuperscript{189}

Approaches to participant protection based on considering members of entire groups vulnerable can have unintended consequences.\textsuperscript{190} Labeling those with specific diagnoses as vulnerable can be potentially stigmatizing by reinforcing gross generalizations about large and varied groups of people.\textsuperscript{191} Some successful initiatives to combat stigma have focused on frameworks of empowerment, which encourage individuals or their loved ones to advocate for ethical policies and practices.\textsuperscript{192} Nevertheless, invoking the concept of vulnerability serves a vital practical and ethical function—it calls our attention to research with human participants that warrants special scrutiny.\textsuperscript{193}

Protecting research participants also requires preventing exploitation—taking unfair advantage of another.\textsuperscript{194} In research, informed consent helps protect against exploitation by providing potential participants with information about the ramifications of participation.\textsuperscript{195} When informed consent cannot serve this function, identifying and establishing alternative or additional protections—especially to prevent clear-cut cases of exploitation—is central to proceeding ethically with research.

The challenge for policymakers and others is to delineate necessary standards for protection while enabling versatile and responsive policies to adapt protections to the needs of research participants in highly variable circumstances.\textsuperscript{196} Scholars have referred to the fundamental tension between under- or overprotection as the “pendulum” of human subjects research protections.\textsuperscript{197} Ethical and regulatory policies formed in reaction to unethical research practices tend to focus on the risks of research rather than the benefits and lean toward excluding potentially vulnerable participants. Many stakeholders in the field see exclusion as overprotective. Over time, policy has
shifted toward maximizing inclusion—a movement that potentially discounts the risks of participation or increases the risk of exploitation.\textsuperscript{198}

Neuroscience researchers must walk a fine line between inclusion and protection. Key to walking this line is ensuring that participants are not being exploited or otherwise being used as mere means to an end. Ethical human subjects research can involve consent by participants or permission from legitimate surrogates to serve as a means to achieve broader societal goals—after all, the goal of research is to produce generalizable knowledge, not to provide therapeutic benefit to study participants. However, treating participants as mere means violates respect for persons.\textsuperscript{199} The ethical safeguards discussed below illustrate some of the steps neuroscientists and others can take to ensure that their treatment of participants with impaired consent capacity is not exploitative. Fully informed consent by the participants themselves might not be possible in certain cases, but participants can still be respected and allowed to participate with adequate protections in place.

\textit{Avoiding and Alleviating Stigma}

When considering and addressing ethical concerns about consent capacity, avoiding policies and practices that perpetuate or exacerbate stigma is crucial. Impaired consent capacity is associated with numerous and diverse health conditions. Researchers often make assumptions concerning an individual’s abilities on the basis of broad generalizations concerning a health condition, thus equating certain diagnoses with a lack of consent capacity. These assumptions are particularly prevalent in social attitudes toward those with mental illness diagnoses.\textsuperscript{200} Such unfounded, often unexamined, beliefs can subject these individuals to stigmatizing and discriminatory practices, in which their individual needs and capacities are not adequately assessed or respected. Stigma can negatively affect an individual’s quality of life through a hindered recovery, loss of legal rights, medical care discrimination, and a shorter life span.\textsuperscript{201}

Several national advisory bodies have recommended policies to protect potential participants with certain conditions that are associated with impaired consent capacity.\textsuperscript{202} However, focusing on groups with certain conditions instead of individuals who might have impaired consent capacity can have unintended, stigmatizing effects. It might imply, for example, that all persons with particular mental illnesses have impaired consent capacity.\textsuperscript{203} Although
certain conditions might be associated with impaired consent capacity, different individuals with the same diagnosed condition can exhibit varying capacities, depending on environment, relationships, context, severity of disease, and neuropsychological functions. Assumptions about individuals’ abilities based on a diagnosis can reflect stereotypes that undermine the respect due to those individuals.\textsuperscript{204}

Ethical neuroscience research can help mitigate stigma and discrimination. For example, neuroscience draws attention to multiple neurological disorders that can impair consent capacity, allowing for a more robust discussion that is not limited to mental illness. Scholars note that effective campaigns to reduce stigma and discrimination must be targeted, local, credible, and continuous.\textsuperscript{205} Efforts that incorporate or originate from the perspectives and experience of individuals affected by such disorders are especially effective.\textsuperscript{206} Neuroscience research on consent capacity itself can identify the underlying neural correlates of cognition and can illuminate how cognitive capacities, including consent capacity, are contingent upon environmental and social cues. By providing a more accurate picture of the diversity and determinants of human abilities, neuroscience might help to undermine common assumptions about what individuals with stigmatized conditions can do and contribute. However, because socially stigmatizing attitudes are not simple to eliminate, education regarding neuroscientific facts alone cannot counter stigma.\textsuperscript{207} Factual information works best when coupled with other strategies, such as increased contact with those who live with stigmatized conditions.\textsuperscript{208}

**History of U.S. Policy Proposals and Recommendations**

Understanding, evaluating, and improving informed consent processes is an ongoing goal of research ethics. The history of research ethics includes multiple efforts by national-level advisory bodies to provide guidance for research involving individuals who might have compromised or impaired consent capacity. This section describes historical efforts from the 1970s to the present, and summarizes previous attempts at proposing regulations or guidance about research involving individuals who might lack consent capacity (see Appendix I: History of Major U.S. Policy Proposals and Recommendations on Consent Capacity in Research for a timeline of the events described here, and relevant others). This history illustrates the challenging tension between
the need for rigorous research on diseases and disorders with high morbidity and mortality, and the need to protect individuals who might be vulnerable because of impaired consent capacity. Federal regulations specific to research involving individuals with impaired consent capacity have never been adopted.

In 1977, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (National Commission) released a report entitled *Psychosurgery*, which outlined ethical protections and limitations on psychosurgery research involving capable individuals as well as those with impaired consent capacity.209 The following year, the National Commission released *Research Involving Those Institutionalized as Mentally Infirm*. That report discussed the need for additional safeguards to protect vulnerable research participants with impaired consent capacity, while recognizing that “prohibiting such research might harm the class of mentally infirm persons as a whole by depriving them of benefits they could have received if the research had proceeded.”210 These reports constituted a response to public revelations about unethical research that had occurred in the preceding decades, often with groups who were institutionalized or who could not provide valid informed consent.

In response to these two reports, the U.S. Department of Health, Education, and Welfare (which later became HHS) drafted proposed regulations to guide research involving adults who lack consent capacity, consistent with the National Commission’s recommendations. Although the reason is unclear, those draft regulations were not adopted.211 Resistance to those recommendations might have stemmed, in part, from their potentially overprotective nature, which could stall potentially valuable and ethical research.

In the late 1990s, the National Bioethics Advisory Commission (NBAC) revisited the topic. In *Research Involving Persons with Mental Disorders That May Affect Decisionmaking Capacity*, NBAC made 21 recommendations in six categories: review bodies; research design; informed consent and capacity; categories of research; surrogate decision making; and education, research, and support—several of which included proposals for new regulations.212 NBAC’s recommendations did not result in new HHS policies or regulations.213 Scholars have speculated about this lack of regulatory uptake. Although NBAC recognized that mental illness should not be equated with impaired
consent capacity, some scholars have criticized its narrow focus on diminished capacity among individuals with mental disorders. Critics have noted that NBAC’s limited focus could stigmatize individuals with mental disorders by implying they are likely to exhibit impaired consent capacity—a concern that NBAC acknowledged in its report. Critics supported a wider focus to ensure the protection of all research participants who might have impaired decision making, regardless of diagnosis. In addition, NBAC recognized the importance of diversity among IRB members, but scholars have raised concerns about the lack of input from a diversity of researchers and patient groups in NBAC’s own deliberations.

Some state legislatures introduced bills modeled after NBAC’s recommendations, but ultimately none were enacted into law. However, in 1999, the National Institutes of Health (NIH) released interim guidance for IRBs that was “generally consistent with the NBAC report.” NIH’s guidance, updated in 2009, provides researchers and IRBs with points to consider when conducting research involving individuals who might have impaired consent capacity. This guidance is intended to help researchers understand how to comply with federal and state regulations and to facilitate consideration of ethical concerns and the dual mission of inclusion and protection.

In 2009, the Secretary’s Advisory Committee on Human Research Protections (SACHRP) addressed research involving individuals who might have impaired consent capacity. SACHRP provided 10 recommendations on several topics, including consent capacity, IRB membership and procedures, participant selection, and LARs. In addition, it proposed a regulatory solution for defining who can serve as an LAR. This guidance reached beyond mental health or psychiatric conditions to address the wide array of conditions that can lead to impairments in consent capacity.

Similar to the proposals that preceded it, SACHRP’s recommendations were not incorporated as official guidance or regulations for researchers. Yet SACHRP’s recommendations are widely cited in current discourse on appropriate additional protections. SACHRP’s delineation of who can serve as an LAR for research is often highlighted, given the lack of relevant legislation in most states.
Current Regulatory Framework

No federal regulations directly address research participation of adults with impaired consent capacity. The Common Rule requires oversight of research by IRBs and requires voluntary informed consent from participants or permission from their LARs. The Common Rule also requires “additional safeguards” when participants might be vulnerable for various reasons, including mental disability, but does not stipulate what these safeguards should be. Protections for certain vulnerable populations exist: Subpart B describes protections for pregnant women and fetuses; Subpart C describes protections for prisoners; and Subpart D describes protections for children. Although previous advisory bodies recommended regulatory changes related to research involving adults with impaired consent capacity, no specific regulations have been promulgated.

Although no specific HHS regulations exist for research involving individuals with impaired consent capacity, as described above, regulations require additional safeguards for vulnerable populations, including “mentally disabled persons.” The HHS Office for Human Research Protections (OHRP) offers some guidance and clarification of how the federal regulations apply to research involving individuals with impaired consent capacity. OHRP notes that the HHS regulations allow an LAR to enroll individuals who cannot provide their own consent into research protocols. In addition, in guidance about research involving participants with potentially impaired consent capacity, OHRP emphasizes that federal regulations require that IRBs possess the necessary professional competence to review research activities, either through IRB members with appropriate experience and expertise or invited consultants. Similar requirements exist in FDA regulations on protection of human subjects. In 2014, FDA released draft guidance on informed consent for research, in which one section addresses research involving participants with impaired consent capacity.
The FDA draft guidance (similar to the NIH guidance document described above) leaves the decision about including individuals who might lack consent capacity to the discretion of IRBs and investigators, and provides several considerations to help address challenges that might arise when enrolling such participants in clinical studies.\textsuperscript{232}

A patchwork of applicable legal protections also exists at the state level. Current laws for designating an LAR to facilitate decisions about medical or clinical care vary by state.\textsuperscript{233} For example, they differ in describing how LARs should make decisions on behalf of patients in the clinical context and who can serve as an LAR. Very few of these state laws address decisions about enrollment in research.\textsuperscript{234}

Within this legal and regulatory framework, uncertainty and lack of clarity remain regarding ethically acceptable research involving participants with potentially impaired consent capacity. A survey of U.S. IRBs revealed considerable variability in IRB policies and practices.\textsuperscript{235} Although some argue that the flexibility and discretion granted to IRBs and researchers by federal regulations are essential for the range of valuable research that can be done, others note that uncertainty about how to protect individuals with impaired consent capacity results in inconsistent practices that are either too restrictive or too permissive.\textsuperscript{236}

### Additional Ethical Safeguards

The Common Rule, FDA regulations and guidance, NIH guidance, and many institutional-level policies and guidance call for additional safeguards for vulnerable research participants, including adults with impaired consent capacity. Respectful and just research policies and practices demand both fair inclusion and additional safeguards for prospective research participants with impaired consent capacity. Relevant safeguards might include assessment of consent capacity, solicitation of assent and respecting dissent, use of independent monitors, potential limits on allowable risk, processes to designate and seek permission of an LAR, research advance directives, and stakeholder engagement.

### Consent Capacity Assessment and Modified Procedures

Respectful policies and practices acknowledge the need to consider consent capacity individually, among diverse participants, and not make blanket capacity determinations applied to all persons with specific diagnoses. Consent
capacity entails the ability to understand pertinent information, appreciate its significance, and use the information to reason and make and express a choice. Consent capacity exists along a continuum and varies among individuals with widely diverse disorders. Investigators should assess consent capacity to avoid making assumptions about prospective participants’ consent capacities on the basis of a diagnosed disorder and thus avoid unfairly labeling and stigmatizing individuals and groups. Robust capacity assessment before research begins (and when indicated, during research) also helps ensure that participants with impaired, fluctuating, or diminishing consent capacity are adequately protected.

Investigators conducting research involving individuals who might have impaired consent capacity should ascertain if potential participants have this capacity. For most clinical research, investigators assess consent capacity informally. Validated assessment tools are available for assessing decisional capacity in both the research and treatment settings. Advantages of these tools include the ability to formalize an otherwise intuitive process at a low cost and to do so relatively quickly. However, the tools vary in content and scoring and have variable validity and reliability. An example of an established tool for assessing consent capacity in research is the MacArthur Competency Assessment Tool for Clinical Research (MacCAT-CR). Researchers can tailor the questions in this instrument to a specific research protocol and obtain a score indicating whether the individual has the capacity to consent.

Regardless of which assessment tool is used, researchers and IRBs should consider whether the participants’ condition, intervention under study, or other environmental factors indicate a need for assessment and reassessment of consent capacity during the course of research. Consent capacity is task-specific and depends on the nature and complexity of the decision at hand. An individual might have consent capacity for certain studies or procedures but not for others. Consent capacity also can fluctuate and might improve or worsen as the individual’s condition changes. For example, with regard to the task-specific standard for capacity, the decision-making capacity of an individual with dementia might be sufficient to meet the ethical and legal standard for a clinical intervention that is therapeutic, but not for an experimental protocol that promises no direct therapeutic benefit. Similarly, a potential participant might have the capacity to consent to a simple and easily understood protocol, but might lack capacity to consent to a more complex
protocol. In addition, decisional capacity can fluctuate, improving or worsening over time, or it can diminish over time as disease progresses, as is often the case with dementia.

An individual’s understanding of the information needed to make a decision depends in part on how the information is presented and explained. Modifying informed consent processes by simplifying forms, orally explaining study procedures, or using creative strategies, such as multimedia supplements, might improve comprehension among participants with certain cognitive or decisional impairments. Understanding more about the underlying causes of impaired decision making could lead to development and testing of effective consent and assessment strategies, such as corrective feedback, repeated explanation, or multimedia consent techniques. Finding effective strategies also could be useful for assessment and consent in the clinical environment.

**Assent and Dissent**

When LARs make research decisions on behalf of individuals who lack consent capacity, researchers should include the individuals in the informed consent process to the extent possible. Many participants lacking consent capacity can still express meaningful desires regarding research procedures, including by indicating assent or dissent. For example, some scholars have proposed seeking participants’ assent and respecting dissent to conduct dementia research in accordance with ethical principles. Respecting dissent serves as a protective measure to avoid inflicting burdens and maintains the dignity of all persons in research. Importantly, seeking assent is not the ethical equivalent of seeking informed consent, because it cannot tell us whether risks to participants are being voluntarily accepted on behalf of others. Nevertheless, meaningful
expressions of assent and dissent are salient, even if insufficient, evidence of participants’ perspectives regarding decisions made on their behalf. Uncertainty regarding the assent and dissent of adults who lack consent capacity persists among both researchers and regulators. The ethical significance of assent and dissent are not equivalent. Assent alone does not provide justification to proceed with research. Although participants can meaningfully indicate assent through a wide range of behaviors, to protect participants from exploitation, passive lack of objection should not be interpreted as assent. Meanwhile, dissent can be expressed as a verbal or nonverbal indication of unwillingness to participate in study procedures and gives researchers a strong reason to refrain from proceeding with research. As an additional form of protection, respect for dissent sets a low threshold for tolerating expressions of discomfort by participants at any time during the course of a study. Uncertainty about the standard of meaningful assent indicates the need for further ethical inquiry, which will help articulate and defend conceptual and ethical standards of meaningful assent and dissent. Such research dovetails with other neuroscience research, which might help delineate evidentiary standards that reflect how we know that certain behaviors correlate to meaningful attitudes.

**Independent Monitors**

Scholars have argued for and against the necessity and desirability of employing independent monitors for research protocols to assess regulatory compliance and such research ethics issues as informed consent. In the context of neuroscience research that might involve participants with impaired consent capacity, or whose capacity is in question, independent monitors can be useful to help researchers consider and address challenges that arise. For example, in some cases, researchers might assume that all participants with a particular disorder lack consent capacity. Independent monitors can facilitate

---

researcher respect for potential participants who might have consent capacity and can exercise their autonomy. In addition, some scholars have suggested that independent monitors can help prevent coercion on the part of researchers who might consciously or subconsciously be motivated by conflicts of interest. Independent monitors who are trained to recognize complexity and nuance can help researchers maintain the highest ethical standards.

Federal regulations authorize IRBs to observe or have an independent third party observe the consent process, including assessment of consent capacity and assent or dissent throughout the research protocol. An independent third-party monitor for consent capacity assessments might be valuable in protocols that have the potential to incur serious risk to participants. Monitors can be trained to observe both verbal and nonverbal cues. They can monitor assent and dissent of participants with impaired consent capacity throughout research and can help determine whether to halt the research with particular individuals on the basis of distress or dissent. Some authors report that independent consent monitors can be a central part of the additional protections in place for certain kinds of research including participants with impaired consent capacity.

**Limits on Acceptable Levels of Risk**

Respect for persons recognizes persons as autonomous and capable of deliberating about personal goals, considering choices and opinions, and determining their own lives. Respect for persons also establishes that “persons with diminished autonomy are entitled to protection,” although some scholars contend that this protection actually stems from the principle of beneficence. In all research, risk to participants must be minimized. Scholars argue that, for research involving participants with impaired consent capacity, research should only move forward if it presents risk below a certain ceiling. The fully informed, autonomous participation of adults in research that poses risk but no prospect of direct benefit furthers important research goals. However, because they have made their own determination about whether participating comports with their values, they are not a *mere* means. By contrast, potential research participants who lack consent capacity are impaired in their ability to make judgments based on their own values and are at risk of being used as mere means. Limiting risk level for participants with impaired consent capacity
can help protect participants from exploitation, by seeking to avoid the most obvious forms of exploitation in which social benefits that result from research are gained regardless of the expense to participants. Limits on risk help prevent some individuals from unknowingly bearing egregious risks of harm for the benefit of others.

The diversity of valuable research involving individuals with potentially impaired consent capacity poses a challenge to uniform recommendations for protecting these participants by limiting risk; research protocols can vary widely in the nature and degree of risk participants might confront. Several advisory bodies have recommended limitations on the level of risk to which adults with impaired consent capacity can be exposed in a particular protocol. Some experts have proposed limiting participation to protocols with the potential for direct therapeutic benefit. Such safeguards are similar to protections in place for research involving other vulnerable groups (such as children) who cannot protect their own interests through informed consent. In the case of research involving children, regulations generally only allow research to proceed if it poses no more than minimal risk or offers a possibility of direct benefit to participants; otherwise, regulations require more stringent safeguards.

Some advisory groups recommend that participants with impaired consent capacity not be included in research without the prospect of direct benefit, unless that research is of “vital importance.” Other groups propose that researchers only recruit participants with impaired consent capacity if the research is relevant to their disorder. Still other advisory bodies have supported the inclusion of participants with impaired consent capacity in minimal risk research, recognizing exclusion as disrespectful and possibly unjust. Less agreement exists regarding the conditions under which it might be acceptable to enroll individuals with impaired consent capacity in research that poses greater than minimal risk and does not offer a prospect of direct benefit. This determination remains at the discretion of individual IRBs.

Legally Authorized Representatives

Participants with impaired consent capacity can be enrolled in certain kinds of research by an LAR. LARs, sometimes referred to as surrogate or proxy
decision makers, are individuals with the legal power to make decisions on behalf of others. State laws dictate who can serve as an LAR, how much decision-making power an LAR has, what kinds of decisions the LAR can make, and what processes and procedures are required to establish an LAR. Using an LAR is one important way to facilitate inclusion of participants with impaired consent capacity in research, ensuring the just distribution of the benefits that might accrue to people who share the disorder under study. Using an LAR also is a reasonable way to help protect participants from exploitation, because loved ones or caregivers who have been designated as LARs (as discussed below) are often the best proxy for representing participant interests.

State laws vary regarding who can serve as an LAR. In most states, health care proxies, or those holding a durable power of attorney for health care previously appointed by individuals when they were capable, are deemed the most appropriate LARs. State laws usually include a list of possible LARs in a hierarchy, including those with health care power of attorney followed by the individual’s next of kin (e.g., spouse, adult children, parents, and siblings). Most state laws describe LARs as having authority for medical decision making, but do not indicate whether the LAR’s decision-making power applies to research participation. Although OHRP guidance indicates that LARs appointed for medical care can make certain research enrollment decisions under applicable state law, uncertainty remains about whether laws specific to medical decisions can or should extend to research decisions. Medical decisions are presumed in most cases to be compatible with the best medical interests of the individual, whereas research enrollment entails procedures or interventions done for reasons other than the individual’s medical interests.

LAR decision making is complex. LARs are often confronted with questions about how to make decisions on behalf of their loved ones. For example, how should they best honor the potential participant’s prior wishes and enduring interests? Should they do what the potential participant would have wanted before their impairment? Or should they attempt to determine what they would want in the present (and how can this be ascertained)? The dilemma is deep: How does even someone as close as a loving spouse or parent ascertain, and act in accordance with, the values, authenticity, and sense of self of their loved ones who no longer have the capacity to consent for themselves?
Current practice and literature encourage LARs to make decisions based on a “substituted judgment” standard. Under this standard, out of respect for the now impaired persons, LARs make decisions based on what the persons themselves would have chosen. This standard encourages LARs to make decisions for other individuals in accordance with those individuals’ preexisting known or presumed values and wishes that are projected forward in time to circumstances under which the individuals no longer have the capacity to consent. However, this is only possible when the individuals’ prior values and wishes are known to some extent. Planning for LAR decision making for clinical care is rare and is even rarer for research decisions.

In making decisions for the impaired individual, LARs consider a series of questions to guide their choices, such as, what did the person value in the past? What do they currently value? Did they make known their wishes about medical treatment or research? How can those values be realized in this context? Vague expressions of wishes regarding research participation can be difficult to translate to the unique circumstances and research protocols that might arise. This decision-making process can be even more complex when it involves individuals who have never had the capacity to form or express values or preferences, or determine what might be their authentic self. When the individual’s earlier values and preferences are unknown or were not developed or articulated in a way that guides a particular decision, LARs often use the “best interests” standard. This standard is based on the principle of beneficence. Under these circumstances, LARs make decisions that are consistent with the individual’s overall best interests.

Evaluation of the research enrollment decision is an additional safeguard when an LAR is making the enrollment decision on behalf of a potential participant. Assessment might include what the LAR understands about the decision they are making, the benefits and risks of the proposed research protocol, motives for enrolling the individual, and their understanding of the values and preferences of the individual for whom they are making decisions. Researchers must remain aware of the possibility of tension between the interests of the LAR and the impaired individual.
Research Advance Directives

In some cases, individuals can prepare an advance directive that specifies their willingness to participate in certain kinds of research before their consent capacity becomes impaired, to be consulted and honored by an LAR. Generally, an advance directive is the designation of a proxy decision maker and a set of written instructions articulated by an individual to direct the actions of others in the future, in case the individual becomes unable to make his or her own decisions. Honoring an individual’s preferences as delineated on an advance directive demonstrates respect for that individual—and it facilitates inclusion of participants with impaired consent capacity while also avoiding exploitation by respecting their preexisting wishes that were intentionally projected into the future for the explicit purpose of consenting (or withholding consent) to health care and research.

One type of advance directive included in the laws of all 50 U.S. states is appointment of a power of attorney for health care, sometimes referred to as a health care agent. Research advance directives, although uncommon, would be especially helpful as part of the informed consent process for research in which the prospective participants’ consent capacity might predictably become impaired at a later date. For example, the NIH Clinical Center’s advance directive for both health care and medical research provides individuals with an opportunity to select broad categories of research in which they would be willing to participate; delineate values, goals, and limitations that should guide their participation in research; and designate a power of attorney to make research decisions.268

However, just as in clinical decision making using an advance directive, practical challenges and ethical concerns associated with research advance directives remain. For example, how closely should they be honored when the wishes of potential participants seem to conflict with the wishes they expressed on paper? Which self should take priority: the person who drafted the directive, or the person with a present-day impairment? A clear-cut way of addressing this challenge might not exist; however, when participants’ current wishes comport with those of their advance directives, the advance directives perform a crucial ethical and practical purpose. They lay as solid a groundwork as possible for respect for persons with currently impaired consent capacities.
The dilemma is greatest when the current wishes of a potential participant conflict with those expressed in an advance directive. But even under these circumstances, we can recognize the practically and ethically difficult question of either to (1) enroll resistant participants in a research protocol against their current will, or (2) enroll willing participants against their advance directive that was carried out under circumstances deemed ethically and legally legitimate for that very purpose.

**Stakeholder Engagement**

Stakeholder and community engagement can help improve informed consent processes, build relationships and trust, and increase the likelihood that research findings are relevant for affected communities. Community engagement is particularly important for research that involves underrepresented and potentially stigmatized groups. *Gray Matters*, Vol. 1 discussed stakeholder engagement as a principal model of ethics integration in neuroscience research. Seeking out the perspectives of persons and groups likely to be involved in research or affected by its results offers the potential to bridge different expectations and understandings of neuroscience research.

Many individuals and groups have a stake in research design, implementation, and results. Increasingly, standard practice in different research areas—especially those with contentious past and present social, political, and ethical implications—is to employ various techniques to identify stakeholders, as well as incorporate and address their perspectives and concerns.

Many approaches to engaging the stakeholders affected by neuroscience research exist, including explicit attention to IRB composition, formal advisory groups, participatory research methods, large public or community meetings, and empirical research designed to elicit stakeholder perspectives. NBAC recommended that IRBs include members or consultants who could contribute understanding of the experiences of those with impaired consent capacity, including current or former patients, family members, patient advocates, or experts on a specific patient population or LAR decision making. Funding agencies, such as NIH, encourage forms of stakeholder engagement in research beyond those that pertain to IRB composition.
Gaps in Our Understanding of Consent Capacity and Additional Protections

The burden of neurological disorders is high and expected to increase considerably as the population ages.\textsuperscript{275} Thus, research that might involve participants with impaired consent capacity likely will continue to increase, as will clinical encounters with patients who have impaired consent capacity. To conduct such research ethically, it is imperative that we learn as much as possible about impairments in consent capacity and how to improve consent processes to protect participants. Our knowledge about conceptual, empirical, and practical matters relevant to consent capacity and what can be done to establish the most inclusive and protective practices is incomplete. Advances in neuroscience can also help advance understanding of cognitive capacity, decision making, and consent. We illustrate some examples of the gaps in our knowledge below, but emphasize that these examples are not exhaustive of the potential for future research and analysis.

\textit{Conceptual Challenge: Defining Consent Capacity and Vulnerability}

Although consent capacity is generally understood to encompass multiple factors, including the ability to understand information, appreciate its significance, use information to reason, and make and express a choice about participation, advances in neuroscience reveal that consent capacity should also account for other aspects that might influence decisions. For example, advances in neuroscience research have prompted scholars to consider whether other aspects of consent capacity such as emotion might have been overlooked. Some researchers report that consent capacity is incompletely understood without examining the emotional aspects of decision making, for example, how mood disorders can affect individuals’ appreciation of risk, including risk aversion and risk tolerance.\textsuperscript{276} Considering emotional aspects of consent capacity might be especially pertinent to neuroscientific interventions known to influence individuals’ mood, such as DBS.\textsuperscript{277}

Scholars sometimes conflate vulnerability caused by impaired consent capacity with other constraints on decision making, such as desperation resulting from lack of treatment options.\textsuperscript{278} Although desperation is distinct from impaired consent capacity, it can affect individuals’ decisions about research participation by altering their risk perception. Other evidence indicates
that concerns about coercion, desperation, or participants’ expectations that research participation confers medical benefit—referred to as the therapeutic misconception—are distinct from whether a person cannot consent because of impaired consent capacity. Continued research to understand the nature of vulnerability and consent capacity could result in better protections for research participants.

**Empirical Challenge: Assessment**

Comparing consent capacity assessment tools is challenging. Such tools have different definitions of reasoning, are tailored to specific protocols, and require various skills and training to administer. Instruments are evaluated for reliability, including both consistency across users and with the same user over time. Empirical data exist on instruments’ validity, sensitivity (ability to detect those with impairments), and specificity (ability to identify those with consent capacity). However, challenges remain in defining consent capacity, enumerating specific skills needed for decision making, and tools for accurately measuring such skills.

Ongoing empirical and conceptual research helps to refine available assessment instruments and is needed to facilitate developing neuroscience. The empirical “gold standard” against which to validate research consent capacity instruments is a psychiatric exam. However, variability exists in psychiatric assessments of capacity. Conceptual challenges also exist. For example, scholars have divergent views about what combination of abilities consent capacity comprises, and identifying required abilities is integral to determining the validity of assessment tools. Researchers must know what abilities to assess to develop good assessment tools. Interdisciplinary expertise is necessary for developing tools that can reliably and accurately measure human abilities in context, as well as determine which score thresholds indicate whether an individual is capable of providing informed consent. These determinations require a variety of expertise, including understanding of the legal and ethical meaning and conditions of consent along with the neurology, psychology, interpersonal and cultural interactions, accommodation, and clinical understanding of the well-specified meaning and conditions of consent. Determining how and when to reassess consent capacity to affirm continued participation requires similar collaboration across empirical and normative disciplines.
Practical Challenge: Developing Standards for Valid Consent and Community Engagement

Some aspects of protections are insufficiently understood because they take place against a backdrop of more general obstacles to ensuring informed consent. For example, even among participants who are capable of providing consent, the therapeutic misconception is thought to be pervasive. Participants also might be subject to a cognitive bias called unrealistic optimism. Researchers are seeking ways to improve information sharing during the consent process and to avoid misunderstandings. Innovative approaches to enhance understanding include shorter and simpler consent forms, and participant engagement to identify what information is best understood, and in what format. Benchmarks for success need to be considered in light of what works in informed consent practices generally.

Similarly unclear is what standards are ethically required for community engagement. Not all neuroscientists will be able to incorporate extensive involvement of community members, for example, because of resource constraints. In other cases, a lack of familiarity with local community leaders and resources might make community engagement difficult. Moreover, even when such engagement is possible, determining what constitutes success is difficult. Such practical challenges have conceptual aspects, including how to determine which individuals or groups constitute a community or who can legitimately represent a community’s views. For example, for certain research that requires community consultation, if or how the research should proceed if a substantial or minority portion of a community express reservations is unclear. For researchers and those responsible for research oversight to know more about which additional protections are ethically necessary or sufficient, researchers must consider both the ethical rationale for certain protections and evaluate what protections they provide in practice.

Recommendations

Consent capacity raises complex ethical questions for neuroscience research. For example, recruitment and retention processes must accommodate variations in consent capacity, and the complexity of enrolling individuals with impaired consent capacity makes participant selection especially challenging.
Some research results might have implications for members of groups whose historical association with impaired consent capacity has subjected them to stigmatization and discrimination or exclusion from research participation.

***

Neuroscientists often conduct research involving participants whose consent capacity might be absent, impaired, fluctuating, or in question, in part because many of the disorders that neuroscience research addresses can affect consent capacity.\(^{291}\) To advance research that seeks to ameliorate these disorders, researchers will need to include affected individuals in studies with ethical safeguards in place. Participants will include affected individuals at any stage of life, from child participants to adults with traumatic brain injury to older participants with dementia.

**Recommendation 6: Responsibly Include Participants with Impaired Consent Capacity in Neuroscience Research**

Researchers should responsibly include individuals with impaired consent capacity who stand to benefit from neuroscience research. Participation, with ethical safeguards in place, can ensure progress aimed at understanding and ameliorating neurological disorders and psychiatric conditions.

Ethical and regulatory protections for research participants, which are influenced by historical revelations of unethical research practices, might lead to overprotection of research participants and exclusion of participants who might lack consent capacity. Contemporary neuroscience offers the potential to better understand—and the hope to one day ameliorate and prevent—devastating neurological disorders and psychiatric conditions. However, to realize this potential and fulfill this hope, affected individuals, including those who might have impaired, fluctuating, or diminishing consent capacity, will need to be included in ethical research with adequate protections in place.\(^{292}\) Responsible inclusion entails compliance with existing regulations, and the use of appropriate additional safeguards, which can vary, depending on the nature of the research and the population being studied. For example, a protocol studying the progression of Alzheimer’s disease might involve continued capacity assessment or an independent monitor who evaluates ongoing consent and capacity assessment processes. An IRB overseeing research
involving participants with impaired consent capacity might choose to limit the acceptable level of risk imposed by the protocol below a certain ceiling.

The pendulum should not now swing in the extreme other direction (as too often happens in history) of under-protection. Rather, we can and should strongly strive toward responsible inclusion. Public beneficence, justice, and respect for persons can ground neuroscientists’ obligation to be as inclusive as possible, consistent with protecting all research participants when designing research protocols. More inclusive practices will ensure that the fruits of neuroscience research reach all individuals who stand to benefit as long as ethical safeguards for all research subjects are squarely in place. In its 2009 report, SACHRP similarly recognized the importance of inclusion to advance scientific discovery and potentially ameliorate suffering.293

Researchers have made substantial progress in the past decade characterizing and understanding consent capacity. However, gaps remain, and further research can support development of best practices for ethical research involving participants with impaired consent capacity. The Bioethics Commission encourages researchers to initiate studies to fill gaps in knowledge and help develop sound policies and procedures for research involving participants with impaired consent capacity.

Recommendation 7: Support Research on Consent Capacity and Ethical Protections

Funders should support research to address knowledge gaps about impaired consent capacity, including the concept of capacity, brain function and decision-making capacity, current policies and practices, and assessment tools.

Conceptual research to address gaps in our knowledge, including the influence of vulnerability, desperation, and affective states on decision making, could lead to better protections for all research participants. Moreover, empirical research evaluating assessment tools and additional protections for participants with impaired consent capacity can determine whether they are adequately protective. Researchers and oversight bodies should develop and evaluate innovative protections for participants with impaired consent capacity. These protections might include novel ways to improve participant comprehension.
and creative research designs that tailor informed consent processes based on information gathered during recruitment, all carefully guided by clear and explicit understandings of the conceptual, ethical, and legal meanings of consent.

Results of such studies should be disseminated widely. For example, professional societies like the International Neuroethics Society or the Society for Neuroscience might encourage members to share ethical strategies at meetings. Investigators also can consider reporting in publications the participant protections that they employed. With the assistance of journal editors, publishing protective consent methods separately or within the manuscript itself—for example, as an “ethics methods” section—could help create a collaborative environment in which neuroscientists facilitate and further best practices in research.

Equating certain conditions with impaired consent capacity or making unfounded assumptions about individual abilities based on diagnoses can exacerbate or perpetuate stigma. In addition to avoiding such pitfalls, ethical neuroscience research also can foster a more accurate understanding of neurological disorders and mental illnesses and potentially mitigate stigma. One principal approach to help neuroscience researchers alleviate stigma is stakeholder engagement.

**Recommendation 8: Engage Stakeholders to Address Stigma Associated with Impaired Consent Capacity**

Funders and researchers should engage stakeholders, including members of affected communities, to build understanding of consent capacity and associated diagnoses to mitigate the potential for stigma and discrimination.

Stakeholder engagement is critical to research design, including identifying what to study and assessing how results might be received. It is also an important safeguard to mitigate potential social harms associated with research participation, cultivate trust, and develop mechanisms to address harms that cannot be anticipated. Stakeholders include those with, or at risk for, impaired consent capacity, caregivers, researchers, and community members affected by research. Stakeholders’ contributions can help mitigate stigma and discrimination by providing information about the lived experience of those
affected by a particular condition. They can reveal the diversity of ways in which a condition might manifest and help dispel common assumptions about certain conditions.

Ethical research involving participants with potentially impaired consent capacity requires that investigators acknowledge the diversity of individual needs, abilities, and relationships with caregivers. Stakeholder engagement provides an additional layer of ethical protection for participants. Researchers gain valuable information about participant or surrogate concerns by engaging stakeholders directly. Stakeholder engagement can help neuroscientists identify and develop practices tailored to specific protocols, disease communities, or categories of impairment. Stakeholder engagement can also guide development of standards for enrolling participants with impaired consent capacity, and it can help determine and address pertinent research questions. During the past two decades, researchers have worked to understand stakeholder perspectives about research participation of persons with impaired consent capacity. Engaging stakeholder communities will help neuroscience researchers uphold ethical standards and craft best practices, remain accountable to the communities with which they work, and foster thoughtful consideration about the potential for stigma and discrimination.

Including affected individuals (those with impaired consent capacity and others) in research is vital to fulfill the promise of neuroscience to ameliorate neurological disorders and psychiatric conditions. The Common Rule requires informed consent from research participants or permission from LARs before research can proceed. Thus, an important step in conducting ethical research involving individuals with impaired consent capacity is determining who can serve as an LAR. Federal and state laws lack clarity about how to make such a determination.

OHRP asserts that, in the absence of a state law authorizing who should serve as an LAR for research, state laws that authorize representatives for medical decision making “may be relevant if the research involves those medical procedures or medical treatment.” However, this nonbinding guidance can leave researchers and IRBs uncertain about who can serve as an LAR. Clarity in identifying LARs will help researchers and IRBs remain accountable
to a clear set of ethical and legal standards for enrolling participants with impaired consent capacity in research.

**Recommendation 9: Establish Clear Requirements for Identifying Legally Authorized Representatives for Research Participation**

State legislatures and federal regulatory bodies should establish clear requirements to identify who can serve as legally authorized representatives for individuals with impaired consent capacity to support their responsible inclusion in research.

Federal bodies can play a role in clarifying how to identify LARs. SACHRP, for example, recommended a list of persons (in order of priority) who can serve as LARs, to be relied upon in the absence of applicable state law. Federal regulatory bodies could endorse SACHRP’s recommendation and explicitly permit researchers and IRBs to rely on SACHRP’s priority list of potential LARs. Alternatively, state legislatures that have not already done so could draft their own priority lists for LARs for research, eliminating the need to rely on lists derived from laws pertaining to medical treatment. Medical treatment laws define consent capacity with reference to medical decisions, not research decisions. They do not address the task-specific nature of capacity to consent to various research protocols. In addition, legislation drafted for a clinical context assumes that the choices offered to LARs are in the patient’s best interest, an assumption that does not necessarily apply to participants in research settings.

---

Research including participants with impaired consent capacity presents challenges to researchers, IRBs, institutions, and regulators. Rapidly advancing neuroscience provides an opportunity to revisit these perennial challenges.

Many federal advisory bodies have offered recommendations and proposed guidance to allow research to move forward while protecting potentially vulnerable participants, but no specific regulations have been promulgated. The Bioethics Commission acknowledges the efforts of these previous bodies and the challenges to implementing new policies. In addition, the Bioethics Commission notes the progress made through research to better understand aspects of consent capacity and urges researchers to continue this work to
further advance our understanding of the conceptual, legal, ethical, and neurological components of consent capacity.

Better evidence and a clearer analytic synthesis—which integrates the many components of understanding a person’s ability to consent in various contexts—are needed to facilitate progress of ethical research that protects participants and seeks to ameliorate the disorders that contribute to impairments in consent capacity. The recommendations outlined in this chapter illustrate four specific areas of improvement that will move neuroscience forward in an ethically responsible manner and which have the potential—if implemented—to pave the way for moving the national conversation beyond impasse. The Bioethics Commission therefore encourages action in light of, rather than in spite of, justifiably grave concerns and remarkable complexity.
CHAPTER 4
Neuroscience and the Legal System
The brains of criminals have captured the public’s imagination for centuries. In 1871, when convicted murderer Edward Rulloff was executed for his crimes, scientists acquired his brain to attempt to discover the neurobiological underpinnings of his wrongdoing. More than 200 years later, in 2013, the Public Broadcasting Service aired the series Brains on Trial, in which actor Alan Alda took the audience through a series of neurotechnologies and techniques, demonstrating how they might be used in an actual criminal trial. Advances in neuroscience offer a better understanding of human behavior, and the potential for improved policymaking, increased accuracy, and decreased errors in advancing justice. The application of neuroscience to the law also raises concerns—some real and some imagined—about scientific reliability, misapplication and overreliance on a developing science, conceptions of free will, mental privacy, and personal liberty.

Neuroscience has a variety of potential applications to the legal system and already is employed in many relevant contexts, including increasingly in criminal law (Figure 2). Prosecutors and defense attorneys use neuroscience evidence in criminal proceedings to support propositions concerning, for example, competency to stand trial, mitigation of criminal responsibility, and predicting future dangerousness. Parties also use neuroscience evidence in the civil context to provide objective evidence of “invisible” injuries, such as toxic exposure, pain, and suffering. Policymakers have invoked neuroscience to advocate for legislation and reform; scholars have advocated using

A sketch of the brain of the convicted murderer Edward Rulloff, drawn by scientists who studied his brain

neuroscience to address biases in legal decision making; and researchers and even some commercial entities have introduced novel uses of neuroscience for investigative purposes.

![Figure 2: Number of U.S. Judicial Opinions Discussing Neurological or Behavioral Genetics Evidence Used by Criminal Defendants, 2005-2012](chart)

A total of 1800 judicial opinions (majority, plurality, concurrence, dissent) issued during 2005–2012 were included. Graph and analysis based on 1586 majority and plurality opinions only. Source: Farahany, N., Database 2014. On file at Duke University.

Neuroscience research holds promise for improving our understanding of human behavior, motivation, intention, and action. However, such research remains in its infancy, and to what extent neuroscience will shape our understanding of these crucial aspects of human behavior is unclear. Moreover, substantial practical limitations constrain what neuroscience is likely to tell us about why particular individuals behave in a specific way; thus, neuroscience might offer greater utility for guiding policy decisions, rather than helping to resolve individual criminal or civil cases.
Nonetheless, neuroscience remains poorly understood by the public, attorneys, and judges. Ensuring the ethical application of neuroscience in the legal and policymaking arena requires substantial public and legal education. The Bioethics Commission seeks to discern what neuroscience can and cannot contribute to our legal system now and in the near future and to facilitate its ethical and scientifically credible use.

Although neuroscience might help us achieve more accuracy in decision making and better policies for trials and sentencing, it does not change the normative or moral questions that the law seeks to answer. Law is a social institution, built on norms developed and instituted by society. Even though neuroscience might guide normative assessments, it cannot solely define them.

**Ethical Analysis**

Use of neuroscience in the legal realm—from the criminal courtroom to legislation and policymaking—warrants ethical analysis. The potential value of neuroscience to improve decision making accuracy and advance justice must be reconciled with the potential for exaggeration, hype, and premature application of scientific evidence and concepts that are not yet validated, well-understood, or interpreted accurately. Ensuring scientific reliability, scientific literacy among decision makers, and engagement by credible neuroscientists and the public will contribute to increasing neuroscience’s value and decreasing unwarranted hype.

**Advancing Justice**

Neuroscience has the potential to advance justice by increasing accuracy in legal decision making and policy development. A deeper understanding of the human brain, cognition, and behavior on both individual and societal levels might help tailor policies and sentences, determine guilt and innocence, evaluate blameworthiness, and predict future behavior. For example, evidence of brain abnormalities might help determine whether a criminal defendant is competent to stand trial. Neurosciencen evidence might contribute to a jury’s determination of guilt or innocence, by helping jurors understand a defendant’s mental state, intent, or voluntariness of action. A deeper understanding of the development and capacity of the adolescent brain might help formulate policies about the sentences that young adults and adolescents should receive.
Neuroscientific techniques like brain imaging might help detect juror bias or determine the reliability of eyewitness testimony. Overall, neuroscience might contribute to more accurate decision making and fairer outcomes. Justice requires that we use empirical evidence, including neuroscience, to strengthen the decisions made in these central civic and political realms.

Enhancing justice by using neuroscience evidence is especially important because of the potentially severe and far-reaching consequences of legal and policy decisions. In the criminal context, punishment can involve deprivation of liberty by imprisonment or the death penalty in some jurisdictions. Such severe consequences warrant particular attention to improving the accuracy of conviction and sentencing. In addition, because legal practice is based on a system of precedent, the application of neuroscience in one courtroom can affect its use in other courtrooms for years to come. In the civil context, courts are clogged with lawsuits that involve disputes over subjective factors like pain and intention, and millions of dollars are spent each year litigating and settling these cases. Neuroscience research efforts to understand and measure these subjective factors can help clarify them and should continue. Interdisciplinary collaboration between neuroscientists and legal decision makers could help advance the cause of justice.

Mitigating Hype

The ability of neuroscience evidence to solve legal and normative questions nearly and cleanly is often exaggerated and hyped. This hype can lead to unwarranted and excessive influence on legal decision makers like judges and jurors. Scientific hype in the media or scientific claims that have not been borne out through replication and verified by the scientific community at large can distort public perception. When legislators rely on hyped scientific claims and unverified science to support political agendas, resulting policies and laws can be unjust.

Thus, neuroscientists, legal decision makers, and scholars must address the tension between advocating the use of neuroscience to improve accuracy and advance justice and prematurely urging its use, potentially hindering justice. The responsibility to avoid hype is shared by many stakeholders, including neuroscientists, members of the media, politicians, judges, and the general public. Public education to improve understanding of neuroscience specifically
and scientific evidence more generally is essential to enhance stakeholders’ understanding of neuroscientific concepts and the limitations of neuroscience within the legal system, and to reduce the potentially negative effects of hype.

**Privacy and Cognitive Liberty**

Accurate, reliable, and relevant neuroscience evidence can and should be introduced into the courtroom and policymaking to advance accuracy and justice. However, in the future, neuroscience evidence might raise concerns about cognitive liberty and invasion of privacy. Some scholars claim that neuroscience brings us one step closer to being able to interrogate the brain or “read minds,” which could have implications for individual privacy.\(^{308}\) However, current technology is extremely limited and is incapable of revealing inner desires, psychological states, or motivations.\(^{309}\) Still, probing the brain through techniques like neuroimaging raises questions about whether inner mental processes deserve more privacy protection than externally observable clues about the mind.\(^{310}\) Even failed efforts to penetrate the mind can offend a sense of privacy. Unlike advancing justice and avoiding hype—two more immediate ethical considerations associated with neuroscience and the law—protecting mental privacy is a forward-looking concern that neuroscientists and legal decision makers might need to evaluate as technology continues to advance.

**Current Use of Neuroscience within the Legal System**

Before further speculating about potential future uses of neuroscience within the legal system and predicting related ethical considerations, the current landscape and the ways neuroscience already guides legal decision making should be assessed. Substantial improvements in neurotechnology and scientists’ understanding of the brain, behavior, and cognition in the past decade have led to an increase in the use of neuroscientific evidence within the courtroom and other legal proceedings.\(^{311}\) Neuroscience use within the legal system follows the recent introduction of behavioral genetics into the courtroom, which presented similar promise to add value and accuracy, but also raised similar ethical and practical concerns.\(^{312}\)

Scholars have argued that neuroscience hype and the fascination with colorful brain images exerts undue influence in legal decision making.\(^{313}\) The question
of whether brain-based evidence has an impact on judges’ and juries’ decisions is an empirical one. Judges and juries have acknowledged the persuasive allure of brain scans, after mitigating defendants’ sentences in light of neuroscientific evidence. In a nationwide survey, almost 200 state trial judges were presented with a hypothetical case in which psychopathy was diagnosed in a convict. About half of the surveyed judges received expert testimony presenting a biological explanation of psychopathy, and the other half did not. Survey results indicated a correlation between inclusion of the biological explanation of neurological disorder with significantly reduced sentence length and an increased number of mitigating factors listed. Additionally, studies have shown that members of the public are more likely to trust diagnosis of a psychiatric condition when brain imaging evidence is presented, as opposed to evidence from psychological assessments. These data indicate that neuroscience evidence can have substantial effects on legal decision making.

**Criminal Law**

Neuroscience has become an integral part of the criminal justice system in the United States. During the past decade, hundreds of criminal cases have been influenced by neurobiological claims. In 2012 alone, over 250 judicial opinions—more than double the number in 2007—cite the use of neuroscience by criminal defendants arguing their brain made them do it. Already, over 5 percent of murder trials and 25 percent of death penalty trials feature criminal defendants using neuroscience to argue for lesser responsibility or punishment.

With a few notable exceptions, scientists are on the sidelines of these legal developments. Researchers often decry use of poorly substantiated cognitive neuroscience and call for caution regarding its use in law.

Many methodological problems stymie the reproducibility and validity of neurological studies of complex behavioral traits. For example, what does “impulsivity” mean? How is it measured? Is aggression on the sports field the same as criminal aggression? How do we measure the difference? How can we disentangle environmental and neurological effects in our advancing understanding of human behavior? Do studies on behavioral variations within a population tell us anything about the causes of specific individuals’ behavior or actions?
Despite these scientific hurdles, the use of neurological evidence in the criminal courtroom is on the rise. Even the gravest of legal decisions—including the question of capital punishment—already have hinged upon neuroscience. Over 1500 judicial opinions issued during 2005–2012 discuss the use of neuroscience by criminal defendants (Figure 3). Almost 40 percent of those opinions pertain to criminal defendants charged with capital murder, and 61 percent involve defendants charged with other serious offenses, including noncapital homicide, assault, robbery, burglary, drug possession, rape, fraud, and theft. Many of these cases include expert evidence about past head or brain trauma, neuropsychological testing, and neuroimaging studies conducted.

Criminal law involves multiple stages of legal decision making, including competency, trial, and sentencing. Neuroscience evidence has been introduced

---

**Figure 3: Claims Made Using Neurological or Behavioral Genetics Evidence in U.S. Capital and Non-Capital Criminal Cases**

<table>
<thead>
<tr>
<th>TYPE OF CLAIM</th>
<th>Capital</th>
<th>Non-Capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitigating evidence</td>
<td>24%</td>
<td>40%</td>
</tr>
<tr>
<td>Competency of defendant</td>
<td>9%</td>
<td>15%</td>
</tr>
<tr>
<td>Mental state of defendant</td>
<td>8%</td>
<td>13%</td>
</tr>
<tr>
<td>Ineffective assistance of counsel</td>
<td>1%</td>
<td>8%</td>
</tr>
<tr>
<td>Other</td>
<td>1.4%</td>
<td>6%</td>
</tr>
<tr>
<td>Claim of legal involuntariness</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>Excuse raised by defendant</td>
<td>2%</td>
<td>5%</td>
</tr>
<tr>
<td>Insanity</td>
<td>3%</td>
<td>5%</td>
</tr>
<tr>
<td>Aggravating evidence</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>Intellectual disability</td>
<td>1%</td>
<td>6%</td>
</tr>
</tbody>
</table>

A total of 1800 judicial opinions (majority, plurality, concurrence, dissent) issued during 2005–2012 were included. Graph and analysis based on 1586 majority and plurality opinions only. Source: Farahany, N., Database 2014. On file at Duke University.
and employed at various stages. Defendants have argued that they should be held less accountable for their actions or punished less severely because of alleged neurological impairments. Prosecutors have seized upon the double-edged potential of such evidence to indict defendants’ characters or underscore future dangerousness. These claims are fueling a reexamination of the criminal justice system.

Competency

The *U.S. Constitution* forbids the trial of a defendant who lacks competency.321 An attorney can challenge a defendant’s competency at any stage of the legal proceedings, from competency to stand trial to competency to be sentenced for a crime. Generally, a defendant is incompetent to stand trial if he suffers from a mental disease or defect that renders him incapable of understanding the charges against him and their potential consequences, or is unable to assist his attorney in his defense.322 Traditionally, psychological evaluations are used to determine competency. Psychologists conduct interviews or behavioral observations to determine whether accused individuals have the capacity to understand the trial or assist in their own defense.

Attorneys are beginning to rely on neuroimaging techniques to supplement psychological evaluation of competency.323 Although perhaps only weakly informative, neuroscience can better ascertain the subjective capacities of criminal defendants than existing tools used in law. Judges typically engage in conversational dialogue with defendants and rely upon their perception of defendants’ responses to assess competency. Neuropsychological testing, a history of neurological trauma, and neuroimaging might improve such judgments. Imaging techniques can determine whether structural or functional abnormalities exist in individuals’ brains that might contribute to a lack of ability to assist at trial.324 Observation of a physical abnormality alone is insufficient to prove incompetence—the individual’s behavior and abilities are relevant. Neuroscience techniques can support competency determinations, especially when fact finders suspect that individuals are lying about their abilities.

Consider for example, David Rothman, a 68-year-old physician, who was charged with conspiring to commit health care fraud against the U.S.
government. Before trial and then before sentencing, Dr. Rothman’s attorney challenged his competency, calling five different experts to testify about the extensive neuropsychological testing and neuroimaging procedures (including magnetic resonance imaging [MRI], positron emission tomography [PET], and electroencephalogram [EEG]) they had performed on him. The experts believed that Dr. Rothman suffered from a severe loss of insight and comprehension relating to a degenerative brain disorder. The state’s counter-expert believed otherwise, based on his limited interactions with Dr. Rothman. After hearing the evidence, the judge agreed with the defense experts, finding Dr. Rothman incompetent to proceed to sentencing. His sentencing was indefinitely suspended.\textsuperscript{325}

**Trial**

Our society, including our legal system, is predicated on understanding and predicting what individuals are thinking, planning, and doing.\textsuperscript{326} Although at first glance, the legal system appears primarily concerned with one’s actions, an individual’s mental state and intention also play a principal role in assigning legal blameworthiness. For example, to be convicted of homicide, one must not only have committed the act of killing someone, but the accused must also have had the requisite \textit{mens rea} (guilty mind)—that is, the intention to have killed the victim.\textsuperscript{327} To make those assessments, criminal law has traditionally relied upon observational methods to assess individuals’ intentions, such as their testimony, the testimony of others who know them, and the observable circumstances surrounding events. In the future, neuroscience might allow us to achieve more accurate and empirical assessments of individuals’ intentions, motives, knowledge, and mental states.\textsuperscript{328} Already, defense attorneys have attempted to use neuroscience to try to prove something about individuals’ mental states, for example, that they lacked the ability to act with purpose.

Some scholars have argued that neuroscience challenges the “folk psychological” beliefs underlying criminal law: that actions are voluntary and the product of conscious choice.\textsuperscript{329} The alternative they propose—that actions arise from unconscious predispositions and decision making over which we have little control—has been largely rejected in criminal law. This is because legal concepts like voluntariness and intentionality are normative and are understood differently by scientists and jurists. This alternative view also does not align well with our subjective experiences of self-directed decision
making. Moreover, research has demonstrated that if we approach individuals as responsible actors, they are more likely to act responsibly.  

To obtain a criminal conviction, prosecutors must prove beyond a reasonable doubt that a defendant acted voluntarily, with the mental state specified by law. Criminal law grants prosecutors a strong presumption that the defendant acted voluntarily. The defense of involuntariness arises only when the defendant’s actions were a reflex or convulsion, or a bodily movement arising from unconsciousness, sleep, hypnosis, or some other factor not in the individual’s conscious control. Circumstances like these are believed infrequent in law. However, scientists and philosophers are currently debating whether neuroscience can show that decision making is primarily unconscious, which, if so, would indicate a fundamental mismatch between legal and scientific understanding of the nature of conscious action.

For example, in a 2009 criminal case, a group of four friends gathered for drinks after work. Later that evening, one of the friends drove the intoxicated defendant’s truck to take them all home. Upon arriving at his house, the defendant found himself locked out of his home, and returned to the truck where he inexplicably assaulted his friend. The three friends quickly regrouped and started to walk away, but not soon enough. The defendant hopped into the truck and drove it into them, injuring one and killing another. At his trial, a neurologist testified that the defendant’s actions were consistent with someone in an automatic, unconscious brain state, taking no purposeful action. The court rejected this claim, finding that the defendant’s actions did not appear unconscious or automatic, but instead met the legal definition of voluntary, intentional, and purposeful. He was sentenced to 15 years to life in prison.

The mismatch between legal definitions of concepts of guilt and neuroscientific understandings of human behavior has stalled attempts to use neuroscience to challenge voluntariness and even mental states in law. Yet, in quite a few criminal cases, defense attorneys have used neuroscience to argue that criminal defendants lacked the mental state to have committed the crimes. Mental state, like voluntariness, has a precise meaning in law: the defendant’s purpose in acting, awareness of the surrounding circumstances, and intent to achieve the resulting consequences.
One problem for defendants in trying to establish their mental state during the commission of the crime is the elapsed time between the crime and later neurological testing. Neuropsychological testing that happens months or years after a crime might have little bearing on the defendant’s brain at the time of the crime. Moreover, although neuroscience can inform defendants’ general behavioral predispositions—such as an inclination toward impulsivity or aggression—it cannot yet tell us defendants’ specific mental states when they engaged in the criminal act.

**NEUROSCIENCE EVIDENCE CAN INFLUENCE JURY DECISIONS**

Peter Jordan Chiesa owned a parcel of land off the highway. Two neighbors with adjoining land accessed their properties by a dirt road that crossed Mr. Chiesa’s land. Over the years, the neighbors clashed regularly about the use, maintenance, and width of the easement. After numerous legal clashes, Mr. Chiesa had the trees pruned on either side of the road. When the neighbors trimmed the trees more, Mr. Chiesa called the police to report that his neighbors were trespassing and indicated that he planned to shoot his neighbors to remove them from his property. He killed two of them.

At trial, the defense argued that Mr. Chiesa—who had no prior criminal record—suffered from brain damage that left him unable to control his emotions and sent him into uncontrollable rages. Computer-assisted tomography (CAT), positron emission tomography (PET), and single-photon emission computerized tomography (SPECT) scans revealed damage in his prefrontal cortex, temporal lobes, and cerebellum—damage that experts claimed would affect his impulse control and temper. His doctors opined that, although Mr. Chiesa was aware of what he was doing when he shot the neighbors, his conduct was driven primarily by impulse, not choice. Despite evident planning—notifying the police of his plan, driving his truck without incident, and aiming the gun at two separate individuals—the jury convicted Mr. Chiesa of the lesser offense of second-degree murder instead of first-degree, premeditated murder.


**Sentencing**

Neuroscience can help society reexamine why and how we punish individuals for committing crimes. Do we do so because the defendant deserves punishment in proportion to the harm caused to society? If so, does a neurobiological understanding of human behavior undermine retributivism as the basis for punishment? Or do we punish individuals who commit crimes to protect society against dangerous criminals? If so, would this goal be better
served by focusing on rehabilitation and reintegration into society for those who commit crimes?

In the majority of criminal cases where neuroscientific evidence has been introduced, it has been used to challenge how severely a defendant should be punished.\textsuperscript{334} This evidence often arises in the context of the separate sentencing hearing that defendants receive when convicted of a murder where the death penalty is being sought, to aid the jury in deciding between recommending the death penalty or life in prison without the possibility of parole.

The results for criminal defendants have been mixed. Even in cases with truly chilling facts, courts have held that an attorney for a capital defendant must investigate a reasonable likelihood of a brain abnormality, or risk being found constitutionally ineffective as counsel.\textsuperscript{335} For juvenile offenders, cognitive neuroscience has strongly influenced the recent constitutional prohibitions on execution or sentencing to life imprisonment without the possibility of parole.\textsuperscript{336} In 2005, the U.S. Supreme Court ruled that persons younger than 18 years of age at the time of the crime could not receive the death penalty, taking note of an amicus brief that argued that juveniles lack fully matured brains, character, and sense of responsibility.\textsuperscript{337} The Supreme Court revisited the question in 2010 and 2012.\textsuperscript{338} Citing favorably from briefs submitted by the American Psychological Association and the American Medical Association, the Court held that life imprisonment without the possibility of parole for nonhomicide crimes committed by adolescents is unconstitutional, and that for homicide crimes, the sentencing must be individualized and consider the maturity of the offender.\textsuperscript{339}

\textbf{Prediction}

Using neuroscience has a potential double edge for criminal defendants. In some instances, neuroscience has enhanced rather than mitigated a defendant’s punishment. Courts have at times regarded defendants’ neurological predispositions as aggravating sentencing factors or circumstances.\textsuperscript{340} If defendants successfully prove that neurological abnormalities contributed to their criminal conduct, courts might regard their brains as \textit{too} broken or \textit{too} dangerous to have at large, even \textit{if} that demonstrates they are somehow less culpable.\textsuperscript{341}
Recidivism prediction (i.e., predicting whether an individual will commit another crime) is used throughout the criminal justice process, including in determinations of bail, sentencing, probation, parole, and treatment program assignments. Methods to predict recidivism include clinical observation and measurement of risk factors, such as age at incarceration, age at release, criminal history, drug use, and social support.

Impulsivity—the persistent lack of restraint and consideration of consequences—is one of the most widely studied risk factors for recidivism. Neuroscience researchers studied the anterior cingulate cortex (ACC), a brain region associated with impulse control, in an experiment attempting to predict future rearrest of previous criminal offenders as a measure of recidivism.\(^{342}\)

Damage to the ACC in humans can lead to changes in inhibition, apathy, and aggression. Although other brain regions also play a role in impulse control, research reveals that the ACC is the most robustly engaged region during impulsive activity.\(^{343}\)

Another study used functional magnetic resonance imaging (fMRI) scans of hundreds of prisoners to ascertain features that distinguish psychopaths from others.\(^{344}\) Studies report that individuals diagnosed with psychopathy compose a large percentage of the prison population and are more likely to reoffend.\(^{345}\) Establishment of biological markers that indicate psychopathy and greater likelihood of recidivism could lead to tailored prison sentences. Some argue that it is unethical to assign longer sentences in response to the same crime for those who are biologically predisposed to reoffending.\(^{346}\) However, others contend that neuroscience might lend more accuracy to existing methods for predicting recidivism, an established practice in the criminal justice system for making decisions about bail, parole, and sentencing.\(^{347}\)

Alternatives to Sentencing

As neuroscience improves our understanding of the brain and neurological correlates of criminal behavior, alternatives to traditional
sentencing and punishment are being proposed and implemented. Several kinds of so-called “treatments” for criminal behavior are already in use, including chemical castration for sex offenders and drugs to combat addiction. Other kinds of regimens are envisioned, such as drugs, psychosurgery, and deep brain stimulation (DBS), to treat impulsivity or aggression.348

“Treating” criminal behavior as an alternative to imprisonment or a condition of release raises concerns about the safety and efficacy of these treatments, and about the coercive nature of such options. For example, Depo-Provera®, the female contraceptive injection that is used to chemically castrate male sex offenders, has not been tested in men—but this off-label use by clinicians is not prohibited by the U.S. Food and Drug Administration (FDA).349 In addition, in some states, chemical castration of sex offenders is mandatory, whereas in others, it is presented as a voluntary alternative to prison or a condition of release. But questions surround whether such a choice can ever be truly voluntary. As neuroscience advances and more treatment options are envisioned—options that might seem like appealing alternatives to a potentially ineffective incarceration system—these outstanding ethical questions must be addressed.350

**Civil Law**

Neuroscience evidence also has been used in the civil law context. Civil law involves suits between private parties that usually involve monetary compensation (as opposed to criminal law cases, which are brought by the government and typically involve such punishment as imprisonment). Workers’ compensation, personal injury, and disability determinations are examples of civil litigation, and they often hinge on establishing “invisible injuries,” such as exposure to toxins or the presence or extent of pain and suffering, making them ripe for the introduction of neuroscience evidence.351 Often, pain can be difficult to establish convincingly in civil cases, and neurological testing and evidence potentially could help evaluate the veracity of claims of pain, including emotional pain or suffering.352

In contract law, neuroscientific evidence has been proposed for intent verification, that is, to measure an individual’s brain activity at the time of contract initiation to determine what the individual intended to agree to and expected to receive in return.353 For example, in one case, a contract for the sale
of land was voided because evidence—including MRI data—demonstrated that one of the contracting parties lacked the mental capacity to enter into contracts. Even in child custody cases, neuroscience has been introduced to establish parents’ fitness or neurological evidence about a child’s development to determine the best interests of the child.

**Accuracy and Errors**

Studies are under way to determine whether neuroscience can be used to improve the accuracy of legal decision making and to decrease errors that arise from bias or limitations of human perception. Research has revealed cognitive biases among jurors and judges. For example, implicit bias by jurors based on race and ethnicity is a common concern in U.S. courts. Implicit bias encompasses attitudes and stereotypes that affect our judgments and decisions but do so unconsciously without an individual’s awareness, intentionality, or control. Because these biases reside deep within our subconscious, they are usually not accessible to individuals through self-reflection. Studies have revealed implicit biases in the courtroom from previously unexamined factors, for example, differences in how severely a judge sentences a defendant on the basis of whether the sentencing occurs before or after the judge has had lunch.

Current strategies to mitigate implicit bias in the courtroom include acknowledging the existence of implicit bias, routinely examining thought processes and decisions to check for possible bias, and increasing exposure to stereotyped group members. Jurors are unlikely to recognize or reveal their own implicit biases. Neuroscience could help by identifying areas of the brain associated with race-related bias and perhaps eventually using neuroimaging techniques to detect juror bias and correct for it. However, concern exists about the feasibility and ethical constraints of using neuroimaging on all jurors. Techniques like this might be most useful in cases where juror bias is being questioned.

Neuroscience also has been used to challenge and reveal limitations in eyewitness testimony—whether through suggestibility that arises from lineup procedures or the limitations of human perception. Ongoing studies are attempting to link brain activation patterns to facial recognition. Using neuroimaging techniques to help witnesses accurately identify offenders is a long way off, but if it becomes possible in the future, it could help to improve accuracy and make just outcomes more likely.
Policymaking

Legislators sometimes rely on neuroscience data and concepts to advocate for their agendas. For science, including neuroscience, to guide just and fair policies and laws, data must be accurate and reliable. Neuroscience might hold greater near-term promise in policymaking than it does in criminal and civil cases. The practical limitations that constrain effective use of neuroscience in the courtroom—including the fact that brain differences across a population do not explain the behavior of any particular individual in the population—might matter less in policymaking than adjudicating trials. To the extent that neuroscience can help us at a group level to understand how and why people behave as they do, it might provide relevant empirical evidence for new and better social policies. However, scientific hype in the media or scientific claims that have not been borne out through replication and verified by the scientific community at large serve only to distort public perception. If legislators rely on hyped scientific claims and unverified science to support political agendas, resulting policies and laws will be unjust.

For example, as described above, the U.S. Supreme Court has relied on neuroscience evidence in decisions that led to substantial changes in juvenile justice around culpability and punishment of adolescent criminals. In developing policy, lawmakers might draw from evidence concerning impaired judgment of adolescents compared with adults, increased emotional reactivity among adolescents, and their increased likelihood to engage in risky behaviors in the presence of peers despite knowing better.360

Scientific evidence should be considered collectively when forming policy decisions. Policy developments based on selected research results or scientific evidence that has not yet been replicated and verified are unlikely to be fair and just. For example, in the 1990s, hyped neuroscience research results led to widespread publicity of a phenomenon called the Mozart effect—an increase in intelligence associated with listening to classical music. The original study demonstrated that college students’ performance on spatial reasoning tests improved for 10 to 15 minutes after listening to a Mozart sonata, which translated into temporarily increased IQ scores during that same period.361 Sustained media interest in the Mozart effect contributed to its false association with children and infants well after the publication of reports
failing to replicate the original results.\textsuperscript{362} Consequently, in 1998, the Georgia state legislature allocated taxpayer money to a policy ensuring that all newborn infants received a recording of classical music.\textsuperscript{363} In this case, unverified scientific results combined with unchecked media hype led to misinformed policymaking and a potential misallocation of taxpayer money.

Disagreement about neuroscience research, findings, and interpretation can have a major impact on the law. Policymakers will likely use neuroscience more frequently as it advances. When mature and verified science is used and interpreted correctly, it can have a positive impact on the law, grounding legislation in evidence. However, neuroscience evidence supporting claims that are not yet fully mature or verified can affect the law prematurely and unjustly.

\textit{Interrogating the Brain}

Courts and police do their best to ascertain the thoughts, intentions, and desires of criminal defendants. Some argue that neuroscience might eventually allow us to be able to interrogate the brain, threatening a deeply held sense of privacy.\textsuperscript{364} But today, and in the foreseeable future, neuroscience does not enable us to read minds. Technology remains extremely limited and cannot reveal the true inner desires, psychological states, or motivations that are worthy of the term “mind-reading.”\textsuperscript{365}

However, novel neuroscience techniques might soon reveal (with a cooperative witness) whether an individual recognizes a face or an object, possesses knowledge relevant to a legal proceeding, is lying or telling the truth, or even allow reconstruction of the visual imagery seen at the time of the crime.\textsuperscript{366} These techniques raise crucial questions, including whether “inner mental or neural processes” deserve more privacy protection than external or behavioral elements such as words and actions, and if so, whether convicting criminals is a sufficient justification for violating that privacy.\textsuperscript{367}

If we could accurately interrogate the brain, with a high degree of reliability, then just as DNA evidence has helped to exonerate many wrongfully accused and convicted individuals, so too might neuroscience offer greater accuracy and insights to improve our laws and policies. We should be open to the possibilities that neuroscience can bring, while ensuring the progress and responsible application of neuroscience to the legal system and policymaking.
Scientists disagree about the accuracy of functional brain imaging technologies at issue here, including what inferences can accurately be drawn from observed brain activity. While researchers debate the meaning and utility of these neuroscience techniques, some commercial entities have already started marketing technologies to interrogate the brain.

From simple consumer-based devices that detect electrical activity in the brain and correlate it with attention and relaxation, to complex fMRI scans to detect brain activity associated with complex thoughts, imagery, and tasks, cognitive neuroscience research is advancing our ability to access and decode basic thought processes in the brain.

Researchers have demonstrated through brain scanning that individuals recognize voices or faces with relatively high accuracy in laboratory settings; however, the results are not accurate enough yet to comfortably rely on this technique in a high-stakes situation. Such research might also demonstrate that a person harbors guilty knowledge—for example, knowledge that only individuals familiar with unreleased details of crimes might have—by measuring physiological responses to recognized stimuli using EEG output. With fMRI and sophisticated computer algorithms, researchers have also been able to roughly reconstruct the visual imagery being observed or imagined by an individual, indicating that perhaps one day society might seek to decode the imagery being visualized by a criminal suspect or an eyewitness.

New lie detection tests, initially commercialized by companies like Cephos Corporation (Tyngsboro, Massachusetts), and No Lie MRI (San Diego, California), use fMRI to analyze the truthfulness of individuals in response to questioning. The premise is that lying is more physiologically taxing than telling the truth, and that measuring blood flow patterns and areas of activation of the brain when answering binary questions can accurately identify lying or truth-telling by individuals. Even though litigants have made several attempts to introduce fMRI-based lie-detection in U.S. court cases, no court has admitted such evidence because of concerns about its scientific reliability. However, some countries are already using similar technologies in criminal investigations. Several reported cases in India involved use of similar technology on witnesses and suspects in criminal cases.
These neuroscience applications raise profound ethical and legal questions. For example, do individuals have a right to mental privacy that safeguards them from being compelled to submit to EEG, fMRI, or other brain-based interrogations? Should eyewitnesses have their memories validated by neuroscience techniques? Do the U.S. Constitution, conventions on human rights, treaties, or other legal and social norms offer adequate protection to individuals concerning these technologies? Does the Fourth Amendment of the U.S. Constitution, which protects individuals from unreasonable searches and seizures by the government, safeguard individuals against such uses? Does the privilege against self-incrimination protected by the Fifth Amendment of the U.S. Constitution protect individuals from being compelled to undergo fMRI-based lie detection? What do we mean by freedom of thought, and to what extent is what we mean consistent with neuroscience findings? As neuroscience and technology move forward, scholars will continue to debate these questions, and the public must keep a close eye on the ethical and societal implications of advances and their applications in the courtroom.

The Value of Neuroscience to the Legal System

It is imperative that we avoid overstatements and hype in discussing how neuroscience can affect the legal system. As one commentator described it, neuroscience will not make the legal system “dry up and blow away.” Measured introduction of neuroscientific evidence and concepts, after they are validated, well-understood, and interpreted accurately, is potentially highly valuable. Neuroscience cannot answer central normative questions that are important to society—for example, why we punish criminals and what it means to be a morally responsible or free human being. However, it can provide us with evidence and support a better understanding of the human brain that might guide society’s consideration of these and other important moral questions. Valid scientific evidence can support evidence-based and more accurate legal and policy decisions. For example, when fingerprinting analysis was first introduced into U.S. courtrooms, it provided evidence relating to whether a defendant committed the act. It does not provide definitive proof; we use other evidence in addition to determine whether the defendant committed the act. Similarly, neuroscience could help address whether an individual has a mental disease or defect—one part of the test of legal insanity in most U.S. jurisdictions. Although neuroscience cannot answer whether an
individual is legally insane—a legal construct to determine which individuals we excuse from blameworthiness—it might get us closer to answering whether an individual satisfies the legal test, or even if that legal test can be improved or modified by a better understanding of human behavior.

As neuroscience progresses, it also can contribute to our understanding of how we reason and make moral decisions. Some scholars argue, for example, that over time, neuroscience will reveal that human behavior can be explained as complex causal chains, rather than self-directed choices arising from free will. They argue that this new understanding of human behavior will not necessarily change the way individuals behave or the way we understand human behavior—people will still have the experience of choosing and of making decisions. However, over time, this neuroscientific understanding could challenge the purpose of punishment in the criminal justice system. These scholars argue that, if moral decisions are just inevitable results of causal chains, then individuals are not blameworthy, and punishment for the sake of retribution is intolerable. In this way, they argue, neuroscience will change the way we think about our legal system. Retributive justice will be replaced by punishment only for the sake of deterrence or opportunities for therapeutic interventions.377

Claims that the law will be revolutionized by new evidence are not novel.378 Even though neuroscience, or any science examining the causes of and contributions to human behavior, can help reveal fallacies in our previously held normative judgments, neuroscience does not inevitably lead us to the answers to fundamental moral questions. For example, even if we can agree that moral decisions result from neurological chain reactions, it does not follow that the individuals who make those decisions should not be held ethically or legally responsible for them. The matter remains of what we as individuals and as a society accept as the standards of ethical and legal responsibility. Ultimately, choices about what law and policy should be are social and ethical choices. These choices can be guided by new discoveries from science, but not answered solely by scientific inquiry.

“Neuroscience, like other research on human behavior and decision making, can help us understand the processes by which we make moral decisions.”

Some scholars have contended that an account of legal responsibility can be defended independently of the existence of free will and that advances in neuroscience will not undermine the basic assumptions of law. Others have used neuroscience to support a legal theory of free will based on freedom of actions. Although science is pertinent to law by providing information about some causes of (and excuses for) behavior, law fundamentally involves normative questions. Law concerns standards for behavior that provide individuals with reasons for action and inaction—either recognition of why an act is wrong or fear of the consequences of violating the law. The determination of legal responsibility concerns not just psychological facts, but social standards or expectations for what it means to have and fulfill legal obligations to other citizens in the community.

Several examples illustrate this point. As described above, neuroscience evidence is used in determining individuals’ competency to stand trial. Previously, psychological evaluations were the only tool at a court’s disposal. Advances in neuroscience now answer questions about competency in a more empirical and potentially more accurate way. The legal elements of competency include an understanding of the charges and an ability to assist in one’s own defense. Neuroscience techniques that might identify the cognitive capacities necessary to understand trial proceedings and assist in one’s own defense could support more accurate competency determinations. For example, neuroscience can address whether an individual has the capacity for long- and short-term memory or whether they have the capacity for sound judgment and reasoning.

Developments in neuroscience likely will continue to refine our understanding of what acting voluntarily or with intentionality means in criminal law. New advances in diagnoses of neurological disorders could provide evidence of involuntary movements, like seizures, which could be used to argue that someone acted involuntarily. Similarly, advances in neuroscience might provide greater clarity about neurological disorders or psychological conditions that cloud consciousness, which might shed light on such symptoms as dissociation and bear on whether behavior was voluntary. Advances in understanding how intoxication and addiction affect psychological states could provide evidence to help jurors differentiate when a defendant has acted recklessly (i.e., with conscious disregard of a substantial and unjustifiable risk of injury to others) rather than negligently, which does not require conscious
awareness of a risk. Neuroscientific evidence might supplement the objective circumstantial evidence being used to determine whether a crime was committed in the heat of passion. Advances in the neuroscience of addiction and mental illness also could be relevant to the legal process of diverting those who have committed crimes into treatment programs, perhaps by revealing evidence of who is most likely to be receptive to treatment.

Advances in neuroscience also could help us better understand how incarceration affects cognitive function and whether it is successful at deterring future crime. Neuroscience research can help support better, more practical, and more humane policies. For example, research indicates that solitary confinement, which involves isolation and restriction of environmental and social stimuli, can cause severe psychiatric harm and negatively affect cognitive functioning. Solitary confinement can cause psychological harm among individuals with no history of mental illness and can exacerbate preexisting disorders. Experiments with animals and humans have revealed the effects of isolation and stimuli deprivation on brain degeneration but, because of challenges in working with prison populations, little neurobiological evidence is available of structural brain differences among prisoners confined to solitary confinement. Research on super-maximum (commonly known as “supermax”) prisons also demonstrates deleterious effects on prisoners’ mental health.

Studies reveal that prisoners are substantially more likely than the general population to have a mental illness or a traumatic brain injury. Self-harm incidents are also more likely among prisoners with a serious mental illness. More research is required on how to treat mentally ill prisoners, how to prevent suicide attempts, and the association between mental illness and recidivism. More empirical neuroscience research will help continue to build understanding of incarceration systems and guide development of better policies. Research about the effects of incarceration on prisoners’ brains and on the prevalence of mental illness in prison also might slowly work to affect society’s views of criminal justice practices and the ethics of incarceration.

Challenges of Applying Neuroscience to the Legal System

Despite the real and potential value of neuroscience application to the legal system, the transition, translation, and interdisciplinary bridge is not without
its challenges, both ethical and practical. Overstatements about the potential value of neuroscience and haste to apply it to the legal system and policy development can lead to misinterpretation and misapplication of scientific information. Drawing legal conclusions on the basis of unreliable science, especially when those conclusions involve the deprivation of an individual’s liberty, represents a miscarriage of justice.

**Practical Challenges**

Although neuroscience can help improve the accuracy of determinations about individuals’ mental states, of note, evidence of defendants’ current mental state does not necessarily pertain to their mental state at the time of an alleged crime. The most relevant judgments about a person’s state of mind are unavoidably retrospective.\(^{394}\) In addition, neurological differences only matter insofar as they correlate to the behaviors of interest. For example, differences in the brains of adolescents and adults do not, on their own, establish that adolescents are less rational. Rather, the association of brain data to behavior, such as irrational behavior, determines the relevance of neuroscientific discoveries to law. In cases such as the adolescent sentencing cases described above, advances in neuroscience supplement or confirm the conclusions that can be drawn from behavioral evidence. Furthermore, the law depends on normative judgments about what it means to be legally responsible. Neuroscience alone cannot determine whether adolescents are reasonable enough to face the death penalty. This determination is not a matter of fact, but a matter of what responses are appropriate for society to employ in light of illegal acts.\(^{395}\)

Although laboratory studies demonstrate varied potential applications of neuroscience discoveries to the legal setting, many of these discoveries are not ready for use in the courtroom. Studies often have very few participants, many drawn from undergraduate student populations—a sample that does not necessarily represent either the population at large or the defendants to whom the studies are intended to apply.\(^{396}\) How to apply research results to individuals in a courtroom setting also is unclear, because scientific findings are typically averaged and statistically grouped.\(^{397}\) Because greater variability can exist between two individuals within a group than between individuals in different groups, determining, for example, which statistical group a defendant
falls into can prove challenging.\textsuperscript{398} This is commonly referred to in the literature as the general to individual (or group to individual) “G2i” problem.\textsuperscript{399}

The G2i problem is a perennial one for courts and represents a fundamental mismatch between population-based science and the individualized determination in law. The goal of science is often to make statistical correlations—for example, neuroscientists typically compile fMRI data from multiple research participants and correlate brain activation patterns with specific cognitive states. But this aggregate information might not accurately reflect an individual’s pattern of activity.\textsuperscript{400} For example, in one study, aggregate data revealed activation of several brain regions during a deception task, but researchers could not accurately predict deception in all individual research participants.\textsuperscript{401} Courts, however, typically are concerned with the individual in front of them rather than aggregate data.\textsuperscript{402} When neuroscientists testify as expert witnesses, they must either extrapolate aggregate data to the individual in question, or they can present group data and allow jurors and judges to make inferences and draw legal conclusions about the individual in question. Each of these options is imperfect, and which option an expert employs varies widely across courts.\textsuperscript{403}

\textit{Ethical Challenges}

Several ethical challenges are involved in the application of neuroscience to the legal system. Some ethical concerns are real and applicable today, on the basis of current uses of neuroscience within the legal system. Others, however, are forward-looking concerns involving questions that might arise in the distant future when neuroscience is further advanced.

One contemporary ethical consideration involves treatment of incarcerated individuals and questions of cognitive liberty. In the United States, executing an individual who is determined to be legally insane is unconstitutional; thus, inmates on death row can challenge their competency for execution.\textsuperscript{404} At issue, and working its way through courts in several jurisdictions, is the question of whether a death row inmate can constitutionally be forcibly medicated to render him competent to be executed.\textsuperscript{405} These questions and the consequences of the court decisions have profound ethical import.
Developments in neuroscience could also provide alternative forms of punishment and rehabilitation for criminals. The prefrontal lobotomy was used for approximately a quarter of a century to mitigate antisocial behavior caused by mental illness. A 2005 study reported that DBS can diminish aggressive behavior; other brain regions associated with criminality might be susceptible to improvement through DBS or even less invasive technologies like transcranial magnetic stimulation. Direct brain interventions in criminal behavior by pharmacological or other neuromodulatory methods might proliferate with advances in neuroscience. As discussed above, this neuroscientific medicalization of crime and punishment is especially concerning, given the vulnerability of incarcerated individuals.

**Recommendations**

Bridging the gap between neuroscience and the law can be difficult. The two disciplines are vastly different from one another, involving different kinds of expertise, assumptions, terminology, norms, and goals. Much of neuroscience aims to make correlations and draw conclusions in the aggregate about human behavior, whereas the law is more concerned with causality and seeks to draw conclusions about individual behavior and motivation. Admitting expert testimony about neuroscience under the rules of evidence can be challenging. In addition, drawing from and interpreting neuroscience in ways that are relevant and probative of important facts, but that do not overly prejudice judges or jurors, is complex. Members of the public, especially ones who will serve as jurors, would benefit immensely from educational resources that help bring high-level neuroscientific concepts into lay terms. Individuals expected to use and interpret neuroscience, including judges and attorneys, would also benefit from greater availability of basic training that helps ease the interdisciplinary transition of neuroscience into the legal decision-making process. Considering the expanding role of the behavioral sciences and neurosciences within the legal system, judges and lawyers must ensure they understand and correctly interpret emerging scientific evidence. Without sufficient training and guidance, lawyers, judges, and jurors cannot effectively assess the evidence and technologies involved in a growing number of legal cases.
Recommendation 10: Expand and Promote Educational Tools to Aid Understanding and Use of Neuroscience within the Legal System

Government bodies and professional organizations, including legal societies and nonprofit organizations, should develop, expand, and promote training resources, primers, and other educational tools that explain the application of neuroscience to the legal system for distribution to members of the public, jurors, judges, attorneys, and others.

Studies reveal the persuasive power of brain scan images and neuroscience information on individual evaluations of the legitimacy of scientific claims. Public education efforts can inform the public and help individuals better understand and interpret scientific claims. For example, Brainfacts.org provides resources designed to educate stakeholders about basic neuroscience principles and societal implications of neuroscience research and its applications. Public education efforts should be responsive to developments in the rapidly changing field of neuroscience.

Specialized training might not prevent novices from succumbing to the allure of neuroscientific explanations, but extended and specific training might be beneficial. Efforts to train lawyers and judges are already underway. For example, the American Association for the Advancement of Science (AAAS) hosts seminars to educate judges on advances in neuroscience and the issues they might encounter as a result of neuroscience developments. Legal professionals demonstrate substantial interest in training sessions. For example, the MacArthur Foundation’s Colloquium for Federal Judges on Law, Neuroscience, and Criminal Justice received 135 applications for just 35 openings in 2013 and 2014. Other resources for lawyers and judges include a neuroscience “boot camp” and several curricular materials published by the MacArthur Foundation Research Network, a Neuro-Literacy 101 course hosted by the Integrative Law Institute, continuing education programs, a forensic psychiatry review course sponsored by the American Academy of Psychiatry and the Law, and a publication funded by the National Center for State Courts, the Open Society Institute, and the State Justice Institute outlining research on implicit racial bias and offering strategies to mitigate it in the courtroom.
In addition to the broad educational tools discussed in Recommendation 10, relevant bodies also should fund and conduct specific research and report results regarding use of neuroscience evidence in making important legal and policy decisions. Organizations and government bodies also should publish reports that address the challenges and limitations of neuroscience’s application to the legal system.

**Recommendation 11: Fund Research on the Intersection of Neuroscience and the Legal System**

Relevant bodies, such as the National Academies of Science, the U.S. Department of Justice, the National Institute of Justice, and the Social Security Administration, should support comprehensive studies of the use of neuroscience in legal decision making and policy development.

In response to the surge in DNA evidence used in criminal investigations and trials and calls from the scientific and legal communities, the National Academies National Research Council’s Committee on DNA Forensic Science produced two detailed reports to address controversy and to answer questions about the use of DNA forensic evidence in criminal cases. These reports proved invaluable to the public and other stakeholders affected by use of DNA evidence. The National Academies similarly issued powerful reports on use of polygraph examinations, recommending against federal government use of polygraphs for screening prospective or current employees to identify spies or other national security risks because of the unreliability of the tests. Similar reports and training materials on the use of neuroscience in law should be developed. These should address the limitations, challenges, and potential for the use of neuroscience in the courtroom and of neuroimaging techniques for investigative purposes.

In addition, untapped resources for empirical work in this area abound, for example, through such government agencies as the Social Security Administration that collect and process data surrounding medical claims in administrative legal proceedings. These agencies might benefit from supporting new studies to aid understanding of the neurobiological bases of pain and disability to facilitate accuracy in claim processing and arbitration.
Neuroscience can add value to legal decision making and policy development—strengthening normative conclusions, lending accuracy and scientific evidence to often intangible information, such as mental state and intent, and potentially weeding out problems of bias and faulty memory. However, neither neuroscience nor any other science will inevitably overturn existing norms undergirding the legal system or fundamentally alter our notions of blameworthiness and responsibility. To maximize the value that neuroscience has to offer, scientists, the media, and legal decision makers must avoid hype. Unrealistically high expectations for new science and technology can lead to a loss of trust when those expectations are unmet.

Recommendation 12: Avoid Hype, Overstatement, and Unfounded Conclusions

Neuroscientists, attorneys, judges, and members of the media should not overstate or rely too heavily on equivocal neuroscientific evidence to draw conclusions about behavior, motivations, intentions, or legal inferences.

As scholars have noted, “[p]eople studying the ethical, legal and social implications of neuroscience have to walk a tightrope.” The ethical implications of potential technologies must be considered before those technologies are used widely. But scholars have been criticized for putting the cart before the horse—puzzling through potential implications of a technology that is not ready for valid and reliable use creates the expectation that it works.

In addition, when neuroscience evidence that is unreliable or has not yet been validated and is not ready for application is introduced into the legal system, justice is threatened. Neuroscientists and courts must exercise caution to maximize the value of neuroscience to the legal decision-making process.

As attorneys introduce more neuroscience evidence into the courtroom, and advocates use neuroscience to influence policy, neuroscientists should engage with the process, consider potential legal applications of their work, and seek to engage with legal and policy decision makers to ease the translation. Neuroscientists can play a principal role in assisting judges and jurors to determine the appropriate interpretations of neuroscientific evidence. Judges need to understand how neurotechnologies work and what their limitations
are, and the limitations of data translation to the courtroom, such as the general to individual problem.\textsuperscript{421}


Neuroscientists should participate in legal decision-making processes and policy development to ensure the accurate interpretation and communication of neuroscience information.

Academic institutions, neuroscience and ethics professional organizations, and science policy organizations can play a role in increasing responsible engagement of neuroscientists with legal decision-making processes and policy development. They can offer educational materials and training resources that describe opportunities for engagement, help neuroscientists understand legal applications for their work, and develop communication skills to bridge language and methodological gaps between the two fields.

Some ways that neuroscientists can engage with these processes include acting as expert witnesses, acting as consultants for a legal team, or helping to file amicus briefs to courts.\textsuperscript{422} Training resources for neuroscientists acting as expert witnesses should emphasize that neuroscientists need to understand certain points, including the different ways that science and the courts approximate truth; that opposing attorneys will attempt to expose flaws in their data and in their testimony; that being an expert witness entails answering specific questions as opposed to delivering a lecture; that judges are responsible for determining whether their expert testimony is admissible; and that legal jargon is different from scientific jargon and that bridging that gap can be challenging.\textsuperscript{423}

\* \* \*

Neuroscience offers a variety of current and potential future applications in legal decision making and policy development. Although neuroscience will not change the fundamental normative inquiries at the heart of the legal system, it provides tools to help address them with more evidence and increased accuracy, and has the potential to advance the cause of justice and the rights of the accused and the incarcerated. Practical challenges to its importation into the legal system persist, and ethical challenges, both real and imagined,
must be considered. Education is essential to combat challenges of hype and reliability and address ethical concerns head-on. Neuroscientists should engage in and lend their expertise to the legal process. The Bioethics Commission urges stakeholders to consider the potential impact of neuroscience on the legal system and policymaking, continue efforts to educate decision makers and others, and strive for reliability and fair and just application.
CONCLUSION
The President charged the Bioethics Commission to examine neuroscience research ethics and the societal implications of applications of neuroscience research as part of the BRAIN Initiative. The Bioethics Commission addressed this charge in two parts. This second part, Gray Matters, Vol. 2, takes an in-depth look at three of the most controversial topics at the intersection of neuroscience and society that have captured the public’s attention—cognitive enhancement and other neural modification, consent capacity, and neuroscience and the legal system. In addition, this volume proposes one overarching recommendation that pertains to all funders associated with the BRAIN Initiative.

Recommendation

In this report, the Bioethics Commission calls for research on a number of critical topics. Such research requires adequate support, including funding, personnel, and other resources. As a White House Grand Challenge, the BRAIN Initiative is uniquely positioned to establish and support efforts that bring together diverse expertise from neuroscience, ethics, law, policy, and other disciplines to advance research and education at the intersection of neuroscience, ethics, and society.

Recommendation 14: Establish and Fund Multidisciplinary Efforts to Support Neuroscience and Ethics Research and Education

The BRAIN Initiative should establish and fund organized, independent, multidisciplinary efforts to support neuroscience and ethics research and education, including the activities recommended in this report.

Financial, practical, and technical support is necessary to investigate important unresolved research and policy questions, many of which are outlined in this report. Organized, independent, multidisciplinary efforts would provide infrastructure to address ethics integration, education, and research. These efforts need not be single brick-and-mortar centers, housed within a particular federal agency or institution. Other configurations can enable stakeholders from an array of public and private entities to collaborate and facilitate research on the ethical and societal implications of neuroscience as the field advances.
Existing initiatives can guide how these efforts might be structured. For example, the HIV [human immunodeficiency virus] Modelling Consortium is a coordinated effort led by representatives from major governmental and nongovernmental organizations. The Consortium identifies questions for further research, facilitates data and information sharing across a diverse set of projects, provides a forum for review of new research, and funds new research. Located around the country, Consortium members meet virtually through periodic meetings and teleconferences. The Human Genome Project Ethical, Legal and Societal Implications Research Program (HGP ELSI) used set-aside funds to support research on related ethical issues. Over the past two decades, a community of interdisciplinary scholars has offered constructive critique of the structure of HGP ELSI and the implementation of its research and education components. The architects of these new neuroscience funded efforts should look to these examples and others to discern successful approaches.

Potential topics that should be addressed by these efforts include the possibility that neuroscience research might contribute to or mitigate stigma and discrimination, be used as evidence in legal proceedings, improve clinical practice, and influence public perceptions of neuroscience and neurological disorders, among others. In addition, these collaborative efforts should support development and dissemination of training materials, guidance, and other educational tools to facilitate informed public debate. Efforts funded by the BRAIN Initiative that integrate neuroscience and ethics will be well-positioned to answer new and remaining ethical questions, consider societal implications of neuroscience research, educate the public, and implement policy recommendations.

***

Advances in contemporary neuroscience research have the potential to deliver both individual and societal benefits, including the potential to treat or cure devastating neurological disorders, and to further our understanding of the human brain and mind. Neuroscience advances have captured the public’s attention and stimulated scholarly and public debate, fueled by accurate accounts of the science as well as hyped and misinformed interpretations. Three controversies represent some of the most important and provocative topics at the intersection of neuroscience and society. Neural modification, including
cognitive enhancement, raises questions about reconciling risks and benefits, ensuring justice, and understanding what it means to be human. Adequately respecting and protecting individuals with impaired consent capacity has presented challenges for decades. Advances in neuroscience and the promise of neuroscience research compel us to reexamine this area—ensuring that those with impaired consent capacity can participate in and benefit from ethical research. Application of neuroscience to legal decision making and policy development offers the potential for more accurate and just outcomes, but also raises concerns about premature use of scientific information, privacy, and moral responsibility. In this report, the Bioethics Commission seeks to clarify the scientific landscape, identify common ground, and recommend ethical paths forward to stimulate and continue critical, well-informed conversations at the intersection of neuroscience and ethics as the field continues to advance.
ENDNOTES


9 *Protection of Human Subjects, Department of Health and Human Services (HHS)*. 45 C.F.R. §§ 46.102, 46.111.


29 Ibid.
32 Topping N. (1952). The United States Public Health Service’s Clinical Center for Medical Research. *JAMA*, 150(6), 541-545, p. 545.
36 McKhann, G.M., supra note 35; Tyler, K., et al., supra note 35.


PCSBI, *supra* note 1.


71 Buchanan, A., supra note 5.

72 WHO, supra note 34.

73 Ibid.


127


Ibid.


Singh, I., Bard, I., and J. Jackson, supra note 97.

Lucke, J., and B. Partridge, supra note 3.
GRAY MATTERS Topics at the Intersection of Neuroscience, Ethics, and Society


102 Outram, S., supra note 89.


106 Farah, M.J., et al., supra note 4, p. 98.

107 Morein-Zamir, S., and B.J. Sahakian, supra note 96; Graf, W., et al., supra note 7.


113 Graf, W., et al., supra note 7.


115 Maslen, H., Faulmüller, N., and J. Savulescu, supra note 70.


117 Savulescu, J., supra note 53.


119 Klonopin® is produced by Genentech, South San Francisco, California.


123 Bostrom, N., and A. Sandberg, supra note 60; British Medical Association, supra note 61.


125 Buchanan, A., supra note 5.


127 Savulescu, J., supra note 53.

128 Caplan, A.L., supra note 121.


130 President’s Council on Bioethics, supra note 54.


132 President’s Council on Bioethics, supra note 54.


134 The President’s Council on Bioethics also argued that people with such memories are essential to provide testimony of atrocities to the rest of society. President’s Council on Bioethics, supra note 54.

135 Ibid; Elliot, C., supra note 133 p. 212.


141 Ibid.

See generally, PCSBI, supra note 139.

Ibid.


146Lucke, J., and B. Partridge, supra note 3.


154 Buchanan, A., supra note 5; Greely, H., et al., supra note 5.

155 Morein-Zamir, S., and B.J. Sahakian, supra note 96; Maslen, H., Faulmüller, N., and J. Savulescu, supra note 70.

156 Maslen, H., Faulmüller, N., and J. Savulescu, supra note 70; Bostrom, N., and A. Sandberg, supra note 60.


159 Graf, W., et al., supra note 7.


162 Greely, H., et al., supra note 5.

163 Graf, W., et al., supra note 7.

164 Ibid.

165 Ibid.


180 PCSBI, *supra* note 143, pp. 24-25.

In addition to the extended time it would have taken to enroll participants had surrogate consent not been used in the NINDS rt-PA Stroke Trial, those enrolled would likely have differed in consequential ways for the generalizability of the results. For example, “subjects enrolled by surrogate consent were older and had more severe strokes than persons enrolled by their own consent. Results from a study using only self-consent would have left clinicians caring for older and sicker patients struggling to determine whether rt-PA was appropriate for this group.” Flaherty, M.L., et al. (2008). How important is surrogate consent for stroke research? *Neurology*, 71(20), 1566-1571, p. 1569.


*Protection of Human Subjects, HHS*. 45 C.F.R. Part 46, Subpart B.

*Protection of Human Subjects, HHS*. 45 C.F.R. Part 46, Subpart A.


Ibid.


133


204 Mental Health Foundation, *supra* note 200.


206 Ibid.


215 Ibid.

216 Ibid.


221 Ibid.


224 Protection of Human Subjects, HHS. 45 C.F.R. § 46.111.
Protection of Human Subjects, HHS. 45 C.F.R. Part 46, Subparts B, C, D.

Protection of Human Subjects, HHS. 45 C.F.R. § 46.111.


Protection of Human Subjects, FDA. 21 C.F.R. § 56.107(a), (f); 21 C.F.R. § 50.20.


Ibid.


See generally, Tovino, S.A., supra note 211.


Dunn, L.B., et al., supra note 238.


Black, B.S., et al, supra note 176. The authors acknowledge that their recommendations might be extended to other domains in clinical neuroscience, including research involving participants who have experienced traumatic brain injury, stroke, or intellectual disability; however, they explicitly caution against using similar standards in a clinical context.


Ibid.


Although some researchers contend that this articulation conflates autonomy and beneficence, disregard for the wellbeing of those who cannot direct the course of their own lives can also be interpreted as disrespectful treatment of valued members of society. Limiting risks of research thus can be said to follow from both the principles of respect for persons and beneficence. National Commission, supra note 166; PCSBI, supra note 186; Beauchamp, T.L. (2005). The Origins and Evolution of the Belmont Report. In J.F. Childress, E.M. Meslin, and H.T. Shapiro (Eds.). Belmont Revisited: Ethical Principles for Research with Human Subjects (pp. 12-25). Washington, DC: Georgetown University Press, pp. 15-16.


National Commission, Research Involving Those Institutionalized as Mentally Infirm, supra note 173, pp. 120-121; NBAC, supra note 202, pp. 51-67; SACHRP, supra note 220.

National Commission, Research Involving Those Institutionalized as Mentally Infirm, supra note 173, p. 120.

Protection of Human Subjects, HHS. 45 C.F.R Part 46, Subpart D; Protection of Human Subjects, FDA. 21 C.F.R. Part 50, Subpart D.

New York State Task Force on Life and the Law, supra note 256.

SACHRP, supra note 220; New York State Task Force on Life and the Law, supra note 256.


270 PCSBI, supra note 1.


273 NBAC, supra note 202, pp. 53-54; SACHRP, supra note 220, pp. 6-7.


275 WHO, supra note 34.


280 Dunn, L.B., et al., supra note 238.

281 Ibid.

282 Sturman, E.D., supra note 237.


SACHRP, *supra* note 220.


Protection of Human Subjects, HHS. 45 C.F.R. § 46.102, 46.111.


OHRP has investigated institutions that have allowed research to proceed using LARs that were not explicitly authorized under state law. See Letter from Kristina C. Borror, Compliance Oversight Coordinator, Division of Compliance Oversight, Office for Human Research Protections (OHRP), to Nathan Kase, Dean (Interim), Mount Sinai School of Medicine. (2002, May 7). RE: Human Research Subject Protections Under Multiple Project Assurance (MPA) M-1155. Retrieved February 2, 2015 from http://www.hhs.gov/ohrp/detrm_letrs/YR02/may02a.pdf.

SACHRP, *supra* note 220.

See generally, Tovino, S.A., *supra* note 211.


310 Ibid.

311 Jones, O.D., et al., supra note 11.


313 Church, D.J., supra note 307.


317 Farahany, N., Database 2014. On file at Duke University. A total of 1800 judicial opinions (majority, plurality, concurrence, dissent) issued during 2005–2012 were included. Graph and analysis based on 1586 majority and plurality opinions only.


320 Farahany, N., supra note 317.

321 U.S. Constitution, amendment VIII; U.S. Constitution, amendment XIV; Pate v. Robinson, 383 U.S. 375, 376 (1966); Dusky v. United States, 362 U.S. 402 (1960) (per curiam). Competence to stand trial and be sentenced is a legal standard that should not be conflated with lay use of the terms “competency” or “capacity.”

322 Offenders with Mental Disease or Defect. 18 U.S.C. § 4241.


324 Elbert, J.M., supra note 303.

ENDNOTES


328 Greely, H.T., supra note 304.


344 Greely, H.T., supra note 304.


Depo-Provera® is produced by Pfizer, New York, New York.


Murphy, E.R., and H.T. Greely, supra note 351.


See, e.g., In re H.P., Np. 07JC19926, 2009-Ohio-2186. (A single mother’s parental rights were terminated, in part, due to her traumatic brain injury and resulting cognitive impairments); In re G.V., 674 S.E.2d 479(N.C. Ct. App. 2009) (Neurological evidence about a child’s development was used to determine whether termination of parental rights was in the best interest of the child).


Casey, P.M., et al., supra note 306.

Werner, N.S., Kühnel, S., and H.J. Markowitsch, supra note 306.


Rosen, J., supra note 308.

Farah, M.J., et al., supra note 309.


374 Greely, H.T., *supra* note 304.


379 Morse, S.J., *supra* note 329.


GRAY MATTERS Topics at the Intersection of Neuroscience, Ethics, and Society

Ibid, p. 162.


Fazel, S., and K. Seewald, supra note 389.


Morse, S.J., and W.T. Newsome, supra note 379.


Ibid, p. 850.

Ibid.


A notable exception is in Supreme Court cases where the Court has carved out categorical exceptions to the applicability of the death penalty for mentally retarded offenders and juveniles. See Farahany, N.A. (2009). *Cruel and Unequal Punishments. Washington University Law Review, 86*, 859-915.


*Morse, S.J., and W.T. Newsome, supra note 379.*


*A notable exception is in Supreme Court cases where the Court has carved out categorical exceptions to the applicability of the death penalty for mentally retarded offenders and juveniles. See Farahany, N.A. (2009). *Cruel and Unequal Punishments. Washington University Law Review, 86*, 859-915.*


*Ford v. Wainwright, 477 U.S. 399 (1986).*

Greely, H.T., supra note 350.

Greely, H.T., supra note 348.


Brainfacts.org is sponsored by the Kavli Foundation, the Gatsby Foundation, and the Society for Neuroscience. The Kavli Foundation, The Gatsby Foundation, and the Society for Neuroscience, supra note 142.


Greely, H.T., supra note 304.

Ibid.

Jones, O.D., et al., supra note 11.

Ibid.

Ibid.

Ibid.

PCSBI, supra note 1.
APPENDICES
Appendix I: History of Major U.S. Policy Proposals and Recommendations on Consent Capacity in Research

1974
DHEW
• **Proposed policy:** Obtain consent of individual or representative with legal authority, establish consent committee
• **Response:** Not adopted

1977
National Commission’s Psychosurgery
• **Recommendation:** Psychosurgery can be performed if national board determines direct benefit, guardian gives consent, court approves
• **Response:** DHEW proposed ban on psychosurgery on those not competent to give informed consent; not adopted

1978
National Commission’s Research Involving Those Institutionalized as Mentally Infirm
• **Recommendation:** Follow Subpart D language; include a consent auditor
• **Response:** DHEW issued proposed regulations

1998
NBAC’s Research Involving Persons with Mental Disorders That May Affect Decisionmaking Capacity
• **Recommendation:** Two levels of risk; need approval/guidelines from national board and consent of legally authorized representative (LAR) if no prospective authorization and greater than minimal risk research
• **Response:** Criticized for having limited scope, being stigmatizing, and as too restrictive; not implemented

1980
National Commission’s Research Involving Those Institutionalized as Mentally Infirm
• **Recommendation:** Follow Subpart D language; include a consent auditor
• **Response:** DHEW issued proposed regulations

1995
NBAC’s Research Involving Persons with Mental Disorders That May Affect Decisionmaking Capacity
• **Recommendation:** Two levels of risk; need approval/guidelines from national board and consent of legally authorized representative (LAR) if no prospective authorization and greater than minimal risk research
• **Response:** Criticized for having limited scope, being stigmatizing, and as too restrictive; not implemented

2002
NHRPAC’s Informed Consent and the Decisionally Impaired
• **Recommendation:** Three levels of risk similar to Subpart D; LAR must give permission; applies to all who lack decisional capacity
• **Response:** Not adopted (charter expired; SACHRP created Oct. 2002)

2009
SACHRP Recommendations from the Subcommittee for the Inclusion of Individuals with Impaired Decision Making in Research
• **Recommendation:** Allow local IRBs to review more than minimal risk research; encourage states to adopt uniform legislation regarding who can serve as an LAR
• **Response:** Not adopted

2001
HHS Working Group on the NBAC Report
• **Recommendation:** NBAC recommendations be applied to all persons with decisional impairment; OHRP and FDA solicit additional public comment
• **Response:** OHRP requested public comment in 2007

2001
HHS Working Group on the NBAC Report
• **Recommendation:** NBAC recommendations be applied to all persons with decisional impairment; OHRP and FDA solicit additional public comment
• **Response:** OHRP requested public comment in 2007

Appendix II: Guest Presenters to the Bioethics Commission Regarding Ethics and Neuroscience

Paul S. Appelbaum, M.D.
Elizabeth K. Dollard Professor of Psychiatry, Medicine & Law; Director, Division of Psychiatry, Law, and Ethics; Director, Center for Research on Ethical, Legal & Social Implications of Psychiatric, Neurologic & Behavioral Genetics, Department of Psychiatry, Columbia University

Giorgio A. Ascoli, Ph.D.
University Professor, Molecular Neuroscience Department; Founding Director, Center for Neural Informatics, Structures, and Plasticity, Krasnow Institute for Advanced Study, George Mason University

Nick Bostrom, Ph.D.
Professor, Faculty of Philosophy; Director, Future of Humanity Institute; Director, Programme on the Impacts of Future Technology, University of Oxford

William D. Casebeer, Ph.D.
(U.S.A.F., Retired)
Program Manager, Defense Advanced Research Projects Agency

Timothy Caulfield, LL.M., F.R.S.C., F.C.A.H.S.
Canada Research Chair in Health Law and Policy, Professor in the Faculty of Law and the School of Public Health, University of Alberta

David Chalmers, Ph.D.
Professor of Philosophy and Co-director of the Center for Mind, Brain, and Consciousness, New York University; Distinguished Professor of Philosophy and Director of the Centre for Consciousness, Australian National University

Anjan Chatterjee, M.D., F.A.A.N.
Professor of Neurology, Center for Cognitive Neuroscience and Center for Functional Neuroimaging, University of Pennsylvania School of Medicine

Mildred Cho, Ph.D.
Associate Director; Professor of Pediatrics, Stanford Center for Biomedical Ethics, Stanford University

Sohini Chowdhury
Senior Vice President, Research Partnerships, Michael J. Fox Foundation for Parkinson’s Research

Miyoung Chun, Ph.D.
Executive Vice President of Science Programs, The Kavli Foundation

Patrick Corrigan, Psy.D.
Distinguished Professor of Psychology, Illinois Institute of Technology

RADM Peter J. Delany, Ph.D., LCSW-C
Director, Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration

Rebecca Dresser, J.D.
Daniel Noyes Kirby Professor of Law; Professor of Ethics in Medicine, Washington University, St. Louis
Margaret Eaton, Pharm.D., J.D.
Former Research Scholar,
Stanford University Center for
Biomedical Ethics; Former Lecturer,
Stanford University Graduate School
of Business and School of Medicine

Martha Farah, Ph.D.
Walter H. Annenberg Professor in
Natural Sciences; Professor of Psychology;
Director, Center for Cognitive
Neuroscience; Director, Center
for Neuroscience and Society;
Senior Fellow, Center for Bioethics
University of Pennsylvania

Howard Feldman, M.D., F.R.C.P.
Professor of Neurology and Executive
Dean, Research; Faculty of Medicine,
University of British Columbia

Erik Fisher, Ph.D.
Associate Director for Integration,
Center for Nanotechnology in Society;
Assistant Professor, School of Politics
and Global Studies and the Consortium
for Science, Policy and Outcomes,
Arizona State University

Paul J. Ford, Ph.D.
Program Director, NeuroEthics Program;
Education Director, Department of
Bioethics, Cleveland Clinic; Associate
Professor, Division of Medicine, Cleveland
Clinic Lerner College of Medicine of
Case Western Reserve University

Hank Greely, J.D.
Deane F. and Kate Edelman Johnson
Professor of Law, Stanford Law School;
Professor (by courtesy) of Genetics,
Stanford Medical School; Director,
Center for Law and the Biosciences;
Director, Stanford Interdisciplinary
Group on Neuroscience and Society
and its Program in Neuroethics, Stanford
Law School; Chair, Steering Committee
of the Center for Biomedical Ethics

Joshua D. Greene, Ph.D.
John and Ruth Hazel Associate
Professor of Psychology; Director, Moral
Cognition Lab, Harvard University

Barbara Herr Harthorn, Ph.D.
Director, NSF Center for Nanotechnology
in Society; Professor, Department of
Anthropology, University of California,
Santa Barbara

Freddie Ann Hoffman, M.D.
Chief Executive Officer and Founding
Member, HeteroGeneity LLC; Former
Deputy Director, Medicine Staff, Office
of the Commissioner, U.S. Food and
Drug Administration

Steven E. Hyman, M.D.
Founding President, International
Neuroethics Society; Director, Stanley
Center for Psychiatric Research, Broad
Institute of Massachusetts Institute of
Technology and Harvard University

Judy Illes, Ph.D.
Professor of Neurology; Canada
Research Chair in Neuroethics; Director,
National Core for Neuroethics; Faculty,
Brain Research Centre, University of
British Columbia
Harry Johns  
President and Chief Executive Officer, Alzheimer’s Association

Deborah G. Johnson, Ph.D., M.Phil., M.A.  
Anne Shirley Carter Olsson Professor of Applied Ethics, Science, Technology, and Society Program, Department of Engineering and Society, School of Engineering and Applied Science, University of Virginia

Christof Koch, Ph.D.  
Chief Scientific Officer, Allen Institute for Brain Science

Walter J. Koroshetz, M.D.  
Deputy Director, National Institute of Neurologic Disease and Stroke, National Institutes of Health

Pat Levitt, Ph.D.  
Chair-Elect, Neuroscience Section, American Association for the Advancement of Science; Provost Professor, Department of Pediatrics, W.M. Keck Chair in Neurogenetics, Keck School of Medicine, University of Southern California; Director, Program in Developmental Neurogenetics, Institute for the Developing Mind, Children’s Hospital Los Angeles

Herbert S. Lin, Ph.D.  
Chief Scientist, Computer Science and Telecommunications Board, National Research Council of the National Academies

Bernard Lo, M.D.  
Professor of Medicine; Director, Program in Medical Ethics, University of California, San Francisco

Peggy Mason, Ph.D.  
Chair, Ethics Committee, Society for Neuroscience; Professor, Department of Neurobiology, University of Chicago

Helen Mayberg, M.D.  
Professor of Psychiatry, Neurology, and Radiology; Dorothy C. Fuqua Chair, Psychiatric Neuroimaging and Therapeutics, Emory University School of Medicine

Robert McGinn, Ph.D.  
Professor of Management Science and Engineering; Professor of Science, Technology, and Society, Stanford University

Deven McGraw, J.D., M.P.H., LL.M.  
Partner, Manatt, Phelps & Phillips, LLP; Former Director, Health Privacy Project, Center for Democracy and Technology

Alfred R. Mele, Ph.D.  
William H. and Lucyle T. Werkmeister Professor of Philosophy, Florida State University

Jerry Menikoff, M.D., J.D.  
Director, Office for Human Research Protections, U.S. Department of Health and Human Services

Jonathan Montgomery, LL.M.  
Chair, Nuffield Council on Bioethics; Professor of Health Care Law, University College London

Stephen Morse, J.D., Ph.D.  
Ferdinand Wakeman Hubbell Professor of Law; Professor of Psychology and Law in Psychiatry, University of Pennsylvania Law School and School of Medicine
Thomas H. Murray, Ph.D.
President Emeritus, The Hastings Center

Ushma Neill, Ph.D.
Director, Office of the President, Memorial Sloan-Kettering Cancer Center; Editor at Large, Journal of Clinical Investigation

Helen Nissenbaum, Ph.D.
Professor, Media, Culture, and Communication, and Computer Science; Director, Information Law Institute, New York University

Carlos Peña, Ph.D., M.S.
Director, Division of Neurological and Physical Medicine Devices, Office of Device Evaluation, Center for Devices and Radiological Health, U.S. Food and Drug Administration

John Perry, Ph.D.
Emeritus Professor of Philosophy, Stanford University; Distinguished Professor of Philosophy, University of California, Riverside

Eric Racine, Ph.D.
Director, Neuroethics Research Unit; Associate Research Professor, Institut de Recherches Cliniques de Montréal; Associate Research Professor, Department of Medicine, Université de Montréal; Adjunct Professor, Department of Medicine and Department of Neurology and Neurosurgery, McGill University

Adrian Raine, Ph.D.
Richard Perry University Professor of Criminology, Psychiatry, and Psychology, University of Pennsylvania

Peter Reiner, V.M.D., Ph.D.
Professor, National Core for Neuroethics, Kinsmen Laboratory of Neurological Research and Brain Research Centre, Department of Psychiatry, University of British Columbia

John Reppas, M.D., Ph.D.
Director of Public Policy, Neurotechnology Industry Organization

Nikolas Rose, Ph.D.
Member, Human Brain Project Social and Ethical Division Steering Committee; Professor of Sociology and Head of Department of Social Science, Health and Medicine, King’s College London

Bruce R. Rosen, M.D., Ph.D.
Professor of Radiology, Harvard Medical School; Director, Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital

Adina Roskies, Ph.D.
Associate Professor, Department of Philosophy, Dartmouth College

Pamela Sankar, Ph.D.
Associate Professor, Department of Medical Ethics and Health Policy; Senior Fellow, Leonard Davis Institute of Health Economics, University of Pennsylvania

Marya Schechtman, Ph.D.
Professor of Philosophy, University of Illinois at Chicago
Terrence J. Sejnowski, Ph.D.
Francis Crick Chair, Professor and Laboratory Head, Computational Neurobiology Laboratory, Salk Institute for Biological Studies; Investigator, Howard Hughes Medical Institute; Distinguished Professor, Section of Neurobiology/Neurosciences, University of California, San Diego

Stefano Semplici, Ph.D.
Chairperson, International Bioethics Committee, United Nations Educational, Scientific, and Cultural Organization; Professor of Social Ethics, University of Rome Tor Vergata

Gregory Simon, M.D., M.P.H.
Investigator, Group Health Research Institute; Chair, Scientific Advisory Board, Depression and Bipolar Support Alliance

Steven L. Small, Ph.D., M.D.
Stanley van den Noort Professor and Chair, Department of Neurology; Professor of Neurobiology and Behavior; Professor of Cognitive Sciences; Director, Neuroscience Imaging Center, University of California, Irvine; Professor Emeritus of Neurology, The University of Chicago

Mildred Z. Solomon, Ed.D.
President and CEO, The Hastings Center; Clinical Professor of Anaesthesia, Harvard Medical School

Nicholas Steneck, Ph.D.
Director, Research Ethics and Integrity Program, Michigan Institute for Clinical and Health Research; Professor Emeritus of History, University of Michigan

Serena Viswanathan, J.D.
Assistant Director, Division of Advertising Practices, Bureau of Consumer Protection, U.S. Federal Trade Commission

Anthony Wagner, Ph.D.
Professor of Psychology and Neuroscience, Stanford University

Stephen J.A. Ward, Ph.D.
Professor and Director, George S. Turnbull Center, School of Journalism and Communication, University of Oregon-Portland

John C. Wingfield, Ph.D.
Assistant Director for the Directorate for Biological Sciences, National Science Foundation

Paul Root Wolpe, Ph.D.
Director, Center for Ethics; Asa Griggs Candler Professor of Bioethics, Emory University

David E. Wright, Ph.D.
Director, Office of Research Integrity, U.S. Department of Health and Human Services

David W. Wright, M.D.
Associate Professor of Emergency Medicine; Director of Emergency Neurosciences, Emory University School of Medicine
Neuroethics in the Age of Brain Projects

Henry T. Greely,1,* Khara M. Ramos,2 and Christine Grady3
1Center for Law and the Biosciences and Stanford Program in Neuroscience and Society, Stanford University, Stanford, CA 94305-8610, USA
2National Institute for Neurological Disease and Stroke
3Department of Bioethics, Clinical Center
NIH, Bethesda, MD 20892, USA
*Correspondence: hgreely@stanford.edu
http://dx.doi.org/10.1016/j.neuron.2016.10.048

Neuroethics advances have brought important ethical questions. The recent launch of two large brain projects, the United States BRAIN Initiative and the European Union Human Brain Project, should accelerate progress in understanding the brain. This article examines neuroethics in those two projects, as well as its exploration by other efforts.

Our knowledge of how brains work continues to advance at a staggering rate. Several recent large, targeted projects—the United States BRAIN Initiative, the European Union’s Human Brain Project (HBP), the China Brain Project, and others—promise to accelerate that rate even further.

Rapidly accruing knowledge about the brain can raise profound questions about who we are, how we function, how those functions might be changed, and what that might mean for society. Understanding how human brains function is perhaps uniquely imbued with ethical, legal, and social implications. Even brain projects focused on developing research tools and largely dealing with non-human brains lay the foundations for possibly revolutionary changes in our human world.

This is a good time to take stock of the state of neuroethics. After describing neuroethics, we will examine first its incorporation into two of the major brain projects, then look at some other organizations promoting research and education in this area.

Neuroethics is a recent area of inquiry concerned primarily with the provenance and implications of knowledge about the brain, and the use of such knowledge to treat, manipulate, and possibly enhance brain function. Many things about neuroethics can be controversial, including even what neuroethics is—a field of its own, a subset of bioethics, or something else. People who publish in neuroethics include ethicists and neuroscientists, but also psychologists, philosophers, lawyers, sociologists, anthropologists, and people from many other disciplines.

Neuroethics can be grouped into three categories: the neuroscience of ethics; the ethics of neuroscience research; and, most frequently, the ethical, legal, and social implications of advances in neuroscience.

Interest in the ethical implications of increasing knowledge about the brain is certainly not new. One can find early concerns about what would now be termed neuroethics in the 1960s and 1970s, particularly in fears about various kinds of “brain control” (Valenstein, 1973). The current era of neuroethics, though, can be dated to about 2002, when several important conferences began to explore ethical issues revealed by new neuroscience findings and techniques, particularly by fMRI.

More recently, we have seen the launch of major projects aimed at furthering the understanding of the human brain. These, too, have precursors. For example, over a quarter of a century ago, in July 1990, United States President George H.W. Bush signed a proclamation, requested by Congress, declaring the 1990s the “Decade of the Brain” (Bush, 1990). But the current surge of interest, a result of more scientific opportunities, comes with more funding—and more support—than the earlier efforts.

**Neuroethics in the BRAIN Initiative**

The United States Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative came a few months before the HBP. President Obama announced the BRAIN Initiative in April 2013 (Obama, 2013a), although the largest part of the initiative was not well defined until after the June 2014 publication of “BRAIN 2025: A Scientific Vision” (BRAIN Working Group, 2014). A working group of the Advisory Committee to the Director of the NIH prepared the report, which set out scientific goals for the project and developed a specific plan for achieving these goals. This plan has guided the NIH component of the BRAIN Initiative and its emphasis on creating and developing tools to accelerate neuroscientific discovery.

Many players are involved in the BRAIN Initiative, including five federal agencies: the NIH, the National Science Foundation, the Defense Advanced Research Projects Agency (DARPA), the Intelligence Advanced Research Projects Activity, and the Food and Drug Administration. Funding has been ramping up quickly since the initiative’s inception; President Obama’s budget for fiscal year 2017 proposes to increase BRAIN Initiative funding to $434 million (White House Office of Science and Technology Policy, 2016). The United States government has been joined in this effort by private firms, universities and research institutes, research foundations, patient advocacy groups, and more, as well as governmental and private partners from Canada, Australia, and Denmark.

BRAIN 2025 lays out seven core principles for maximizing the value of the BRAIN Initiative (as well as a set of seven priority research areas). The sixth principle was to “consider ethical implications of neuroscience research.” The report said that “the research supported by and the knowledge generated through...”
the BRAIN Initiative should be regularly assessed for their ethical, legal, and societal implications.” It added, as one of the ethics deliverables, “Vigorous dialogue among ethicists, educators, government and corporate representatives, patients and their advocates, lawyers, journalists, scientists and other concerned stakeholders about social and ethical issues raised by new knowledge and technologies generated under the BRAIN Initiative” (BRAIN Working Group, 2014).

The report then laid out four explicit goals for the ethics component for years one through ten of the Initiative:

1. Joint neuroscience/ethics training programs and meetings to consider the unique issues raised by human neuroscience research, and to establish a shared vision for the ethical conduct of such research.

2. Resources for collecting and disseminating best practices in the conduct of ethical scientific research, particularly for the conduct of clinical research.

3. Support for data-driven research to inform ethical issues arising from BRAIN Initiative research, ideally with integrated activities between ethicists and neuroscientists.

4. Opportunities for outreach activities focused on engaging government leaders, corporate leaders, journalists, patients and their advocates, educators, and legal practitioners in discussion of the social and ethical implications of neuroscience research (BRAIN Working Group, 2014).

Shortly after the publication of BRAIN 2025, NIH created a “multi-council working group” (MCWG) made up of non-governmental representatives from the advisory councils of each of the ten NIH institutes or centers that contribute to the initiative, as well as five at-large members (NIH, 2016a). MCWG has met twice a year since August 2014 to provide information and perspectives to NIH about the scientific vision of the BRAIN Initiative in the context of the evolving neuroscience landscape. The MCWG meetings include representatives from other federal agencies, foundations, and international groups partnering with the Initiative.

Meanwhile, on July 1, 2013, President Obama asked his Presidential Commission for the Study of Bioethical Issues (PCSBI) to engage with the scientific community and other stakeholders, including the general public, to identify proactively a set of core ethical standards—both to guide neuroscience research and to address some of the ethical dilemmas that may be raised by the application of neuroscience research findings. (Obama 2013b)

The commission’s undertaking and response, concurrent with the development of “BRAIN 2025,” culminated with the publication of a two-volume report called Gray Matters. The first volume, released in May 2014, recommended integrating ethics early and explicitly throughout neuroscience, articulated the goals of such integration, and described several existing models that could be adapted or expanded (PCSBI, 2014). The second volume of Gray Matters, released in March 2015, recognized that neuroethics is a rapidly growing, multidisciplinary field encompassing a diversity of ethical issues, some of which are not unique to neuroscience but may be expressed in sharper relief, and some of which have the potential to affect the essence of who we are as humans (PCSBI 2015). It focused on three controversial topics in neuroethics: cognitive enhancement, consent capacity, and neuroscience and the legal system.

But neuroethics in the BRAIN Initiative needed close and continuing attention, which the commission was not well placed to provide. As a preliminary step, the NIH held a one-time workshop on neuroethics in November 2014 (NIH, 2014), and the following August, MCWG approved the creation of a subgroup, a Neuroethics Work Group, drawn from MCWG membership as well as other experts in neuroethics (NIH, 2016b; Greely, 2016).

The work group currently has nine members, all unpaid volunteers, and is co-chaired by two of us (C.G. and H.T.G.); the third author (K.M.R.) is the work group’s executive secretary. The work group provides input to MCWG, and through it to the NIH, about ethical issues in its proposed areas of work, occasional consultations on particular research projects, and general information to investigators on spotting, and dealing with, ethical issues. It is also collaborating on workshops around the country on ethical, legal, and social issues the BRAIN Initiative raises. In addition, the work group helped inform the NIH planning for an early 2016 offer to fund supplements to BRAIN researchers for neuroethics issues: its Request for Information on Neuroethics, released in May 2016 (NIH, 2016c), and its intended Request for Applications for funded research projects in neuroethics, approved as a concept by MCWG in July 2016 and released on October 21, 2016 (NIH, 2016d).

At least one other federal agency that is part of the BRAIN Initiative has taken a similar step. DARPA has a seven member Neuroethics, Legal, and Social Issues Advisory Panel. The panel meets at least quarterly to address specific issues that arise in connection with ongoing projects in the neurotechnology work of DARPA’s Biological Technologies office.

The work group is new and still very much a work in progress, but it is a major NIH effort to deal appropriately with possible ethical issues raised by the BRAIN Initiative. But what are those issues?

Most are issues that arise fairly directly from the research supported by the BRAIN Initiative. At NIH, that research focuses heavily on creating tools that speed up the generation of medically oriented research. It also generally has involved research with non-human animals. More recently, issues arising from implanting devices into human brains have become a major focus. What consent requirements are appropriate? What are the long-term responsibilities of researchers (and funders) to people who received the implanted devices, particularly those who had good results? Issues both similar and different will arise from non-implanted, non-invasive methods of recording or modulating human brain activity. Privacy and confidentiality are additional important points. The collection and sharing of information on brains is seen as
crucial to effective research but might go beyond what the human sources of that information expected and may even put their privacy at risk. Although many issues, such as privacy, safety, fairness, freedom and free will, personal identity, and moral responsibility are familiar issues for bioethics, they often take on intriguing new dimensions and complexities because of the power of science and technology and what the brain represents.

As the BRAIN Initiative proceeds, it will engage in more and more research directly with people and directly relevant to human society. As it does, the ethical issues it will confront will necessarily change and expand, as will the efforts of the Neuroethics Work Group.

**Neuroethics in the HBP**

In October 2013, the European Commission started the HBP, one of two winning applications for a new research model, the Future and Emerging Technologies flagship (Human Brain Project, 2016a). Under this model, the European Union funds a core project, which in turn supports and interacts with various partner projects, funded nationally. The HBP consortium of researchers includes more than 115 partners in 24 countries and an expected 10 year budget—from the European Union, national governments, and private partners—of €1.2 billion. It is coordinated by the École Polytechnique Fédérale de Lausanne and is based in Geneva.

Funding for HBP is to be provided in phases. The Ramp-Up Phase ran from October 2013 to March 2016. In 2014, hundreds of European neuroscientists criticized the HBP in an open letter, expressing concern about the project’s focus on databases and computational structures leading to simulating human brain function. Management changes were made in 2015, and the HBP has now moved into Specific Grant Agreement Phase 1, which runs from April 2016 to April 2018. The European Union will provide €89 million to fund this phase.

The HBP is governed by a Stakeholder Board, made up of representatives from countries that make major financial contributions to the project. The Stakeholder Board operates through a directorate that in turn works through a Science and Infrastructure Board and a Project Coordination Office. The work ultimately is done by twelve subprojects. Five of them focus on particular research topics, such as human brain organization; six of them are platforms that are to construct the HBP research infrastructure, such as the neuroinformatics platform; and one, Central Services, coordinates the project.

Neuroethics has a definite place in the HBP (Human Brain Project, 2016b). One of the five subprojects focusing on particular research topics is Ethics and Society. This is charged with performing research on the ethical and social effects of the HBP’s work. It has employees charged with managing the project’s ethics activities and is administered by the Ethics Director (called the Ethics Manager in the Ramp-Up Phase). It awards grants to outside ethics researchers as part of the subproject. It also gets information from designated rapporteurs in the other research subgroups of the HBP.

The Ethics and Society subproject has, thus far, focused heavily on issues of privacy and on research involving non-human animals, particularly on non-human primates. It is also responsible for trying to coordinate as necessary the ethics standards of the many different countries involved in the HBP, sometimes through writing standard operating procedures for the ethical conduct of the project’s research. One of its research grants went to Professor Nikolas Rose and the Foresight Lab at King’s College London to identify and evaluate potential ethical and social effects, in part through discussions with researchers in the HBP (Rose, 2014).

In addition to the Ethics and Society subproject, the HBP has an Ethics Advisory Board (EAB). This board was formed by the merger in September 2015 of two initial external advisory boards, one specifically on research ethics and the second on the broader ethical, legal, and social aspects of the project. The new board has eleven members, who, except for the chair, are, like members of the BRAIN Initiative Neuroethics Work Group, unpaid. It is to provide advice to the HBP’s board, both at the board’s request and on its own initiative. As a paid member of the HBP, the Ethics Director serves as a liaison between the Ethics Advisory Board and the Ethics Subproject researchers.

Additional components of the HBP ethics efforts include the Ethics Rapporteur system, which has a representative from each research subproject work with EAB members to identify key ethical concerns within their subprojects. The HBP Ethics group has also established a “Point of Registration” for researchers or members from the external community to register ethics concerns, which are managed by the EAB and Ethics Director. Finally, in April 2016, an independent Ethics Ombudsperson, not paid by the HBP, was established to allow anonymous reporting of concerns about the HBP.

**Comparisons and Contrasts**

Both projects have ethics components, similar in some ways and different in others. They are alike in that the ethics processes in both are still very much works in progress. Both groups are focusing primarily on the direct ethical issues raised by the research they are funding, though the remit of the HBP may reach farther. Both are providing grants for ethics research as well as doing some ethics work directly. The HBP does have employees paid to work on ethics issues through its Subgroup 12, which the Neuroethics Work Group lacks. Approximately 4% of the HBP budget goes to ethics research and to managing ethical issues that may arise.

The Neuroethics Work Group comes largely from the American bioethics tradition as influenced by work on Ethical, Legal, and Social Implications (ELSI) of new biotechnologies, particularly genetics. The HBP seems to be more influenced by an approach called Responsible Research and Innovation (RRI), which grows more from sociology and science, technology, and society approaches than from bioethics.

The two projects are beginning to interact with each other in the mutual hope of advancing the common ethical and social interests of their two projects.

**Neuroethics beyond the Big Brain Projects**

This article’s discussion of big brain projects has focused on the BRAIN Initiative and the HBP. Other countries have announced, and funded, research efforts around the brain, including Australia,
Canada, China, Cuba, Japan, Korea, Sweden, and others. Some of them are also pursuing neuroethics issues. But more importantly, neuroethics goes far beyond just brain projects, both in issues and in organizations.

The brain projects, especially the BRAIN Initiative but also to some extent the HBP, understandably have focused on the ethics of the research they are supporting or, sometimes, their near term consequences. For instance, NIH’s mission focuses its interests squarely in the biomedical domain. The ethics of brain research are one important component of neuroethics, but only one. The projects have not tried to touch another major aspect of neuroethics, the neuroscience of ethics. And more importantly, they have so far done little to investigate the likely effects of neuroscience advances on society. (The HBP may be supporting more research in this than the BRAIN Initiative.)

For example, lawyers and judges as well as scholars have viewed the implications of neuroscience for the law as highly important. Over a thousand articles and book chapters have been published on these topics in the last decade (MacArthur Foundation Research Network on Law and Neuroscience, 2016a). These have often looked at issues of responsibility (which has also been examined in nonlegal, moral contexts), but they have also worried about more concrete issues, like using neuroscience methods to determine a person’s competency or a witness’s honesty. Today, these kinds of questions are often both too far downstream and beyond the scope of the brain projects, but that may change.

Similarly, scholars have worried about the ethical, legal, and social consequences of neuroscience-based cognitive enhancement, memory manipulation, mind reading, marketing, and disease prediction. None of these seems likely to be addressed any time soon by the ethics components of the brain projects.

Happily, other complementary efforts have been, and continue to be, made. The International Neuroethics Society is one of the most important (International Neuroethics Society, 2016). Founded in 2006, it is a scholarly society for students, trainees, teachers, researchers, and members of the general public who are interested in these issues. It holds an annual meeting, helps sponsor other conferences and events, and provides background information to the media and advice on policy issues. Fundamentally, though, it promotes contact between people working on neuroethics in order to stimulate further research and teaching on the subject.

The MacArthur Foundation Law and Neuroscience Network has been another important participant in neuroethics discussions (MacArthur Foundation Research Network on Law and Neuroscience, 2016b). Established in 2007, it has gone through two phases, both focusing on criminal law, with the second phase adding a strong emphasis on juvenile misconduct. It has been a leader among many organizations, such as the Federal Judicial Center, the American Association for the Advancement of Science, the National Center for State Courts, and the American Bar Association, that have provided education about neuroscience to judges. The MacArthur Network is reaching the end of its life, but it will leave behind a wealth of research and resources, including a comprehensive bibliography of work on law and neuroscience.

Another, very new organization is Our Brain, Our World, Ourselves (or O3). It is a global consortium that aims to build a neuroethics alliance that incorporates individual funders, ethics expertise, and public stakeholders to advance neuroethics as a global concern with an emphasis on involving the entire world, particularly countries that have not been in the forefront of neuroscience research (Our Brain, Our World, Ourselves, 2016).

Many other organizations have contributed to work in neuroethics. Sometimes these have been neuroscience groups; the annual meeting of the Society for Neuroscience has long featured both a plenary talk on ethical issues as well as neuroethics-focused events. Broader groups, such as the American Association for the Advancement of Science; the United Kingdom’s Royal Society; and the United States National Academy of Sciences, Engineering, and Medicine, have supported projects, conferences, workshops, and other explorations of neuroethics. Two dedicated publications have also supported work on neuroethics. Starting in 2007, The American Journal on Bioethics published two issues a year labeled “AJOB: Neuroscience” and focusing exclusively on neuroscience, although it had sporadically published occasional articles on neuroethics earlier. At the beginning of 2010, these issues became an independent publication, American Journal of Bioethics: Neuroscience, which has been published four times a year since then. In the meantime, March 2008 saw the launch of the journal Neuroethics, which publishes three issues a year. It is fair to say that the American Journal of Bioethics: Neuroscience has had a more North American focus, while Neuroethics has tended to publish more European scholars, although both journals have published articles from authors around the world.

These two dedicated journals have not been the only homes for neuroethics publications. Scientific, medical, philosophy, legal, and bioethics journals have participated in the increasingly large and interesting explorations of neuroethics issues. This journal itself published one of the first important articles in the current era of neuroethics in 2002 (Roskies, 2002). But in whatever outlet, neuroethics research and scholarship have been expanding rapidly.

Conclusion
Neuroscience is thriving, increasing our knowledge of how brains work at an accelerating rate, in part as a result of the BRAIN Initiative, the HBP, and other large projects, as well as the vast amount of research undertaken by individual investigators, private non-profit organizations, smaller brain-oriented projects, and commercial firms. Knowledge is exploding in many areas of science, but none is likely to have as great an ethical, legal, and social effect as neuroscience. Cosmology and computation are wonderful, but they are not about “us.” Neuroscience, like all of biology, is about “us,” but, even more, in what it teaches us about human brains and consequent human behavior, functions, and consciousness, it is arguably about the most important aspect of “us.” As neuroscience grows, so grow its ethical, legal, and social implications…and so must grow neuroethics.
ABOUT THE AUTHORS

Henry T. Greely is Professor of Law at Stanford University, where he directs the Stanford Program in Neuroscience and Society. He is one of two co-chairs of the Neuroethics Work Group of the Multi-Council Working Group for the NIH BRAIN Initiative. His research focuses on ethical, legal, and social issues arising from the biosciences. Khara M. Ramos is a Senior Science Policy Analyst in the Office of Scientific Liaison within the Office of the Director of the National Institute of Neurological Disorders and Stroke, a part of the NIH. A neuroscientist and former AAAS Science and Technology Policy Fellow, she has broad interest in neuroscience research, with specific focus on how advances in neuroscience intersect with society. Christine Grady is the Chief of the Department of Bioethics at the NIH Clinical Center, a member of the President’s Commission for the Study of Bioethical Issues, and the other co-chair of the Neuroethics Work Group. Her research focuses primarily on the ethics of clinical research.

ACKNOWLEDGMENTS

The authors particularly want to thank Dr. Karen S. Rommelfanger for her assistance.

REFERENCES


BRAIN EQUALITY: LEGAL IMPLICATIONS OF
NEURODIVERSITY IN A COMPARATIVE
PERSPECTIVE

Andrea Lollini*

This article inquires into some of the implications of the concept of neurodiversity. First, it analyzes the definition of neurodiversity and its legal dimension. Then, it explores the claim made by part of the neurodiversity movement that people with different neurodevelopment profiles should be considered a new minority. Finally, this article discusses how neurodiversity might require a new interpretation of the idea of constitutional equality. In order to discuss these issues, this article comparatively analyzes neurodiversity-related jurisprudence of the highest North American and European courts over the last fifteen years. Examining these decisions helps determine the current relationship between neurodiverse individuals and legal systems, highlighting that the principle of equality is under intense pressure when the context implicates neurological diversities.

I. INTRODUCTION ................. 70
II. NEURODIVERSITY: DEFINITIONS, AND DEBATES ..... 74
III. LEGAL NEURODIVERSITY .......... 78
   A. The Size of the Neurodiverse Population ..... 79
   B. Legal Implications of Neurodiversity ............. 83
IV. NEURODIVERSE INDIVIDUALS AS A NEW MINORITY? .... 91
V. HIGHEST COURTS AND NEURODIVERSITY-RELATED
   LITIGATION IN A COMPARATIVE PERSPECTIVE .... 96
   A. Education, Health Care, and Social Services .... 97
   B. Physical Liberty, Competency, and Criminal
      Law ........................................ 115

* This article is the result of a research project: NEDBELS (Neurodiversity Between Law and Science). NEDBELS has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No 656988 (Individual Global Fellowship). NEDBELS was a collaboration between UCSF-UC Hastings Consortium on Law, Science and Health Policy and the University of Bologna (Dipartimento di Scienze Politiche e Sociali), in partnership with UCSF The Alba Language Neurobiology Lab, the UCSF Star Center for ASD and NDDs, and the Istituto di Scienze Neurologiche di Bologna. NEDBELS is deeply grateful to David Faigman, Maria Luisa Gorno Tempini, Bennett Leventhal, Paola Visconti, Antoine Garapon, Jaime King, Ugo Mattei, Sarah Hooper, Dorit Reiss, Robert Schwartz, Radhika Rao, Tatiana Pollard, Francesco Palermo, Roberto Toniatti, Roberto Bin, Marco Cammelli, Daniela Piana, Giovanna Endrici, Silvia Cei.
Can atypical neurocognitive traits be considered new constitutional grounds of discrimination? Are brain-based diversities the new frontier of unfair treatment and injustice? What if we ascertained, with the advancement of neuroscience, that a very large portion of the human population interacts socially, de-codifies body communication, uses body-language, processes information and stimuli, learns, rationalizes, or makes abstractions with a greater variation than previously postulated? What are the implications of the fact that this diversity in human cognition is often considered a clinical syndrome? What if we assumed that unfriendly, uncaring, or biased social environments could make those diverse neurocognitive traits extremely maladaptive? What if, in the end, we discovered that in our current state we lose a considerable human potential—while simultaneously raising the cost of managing those who do not fit what society considers normal patterns of cognition? In other words, are brain attributes one of the cornerstones upon which inequality and injustice are built in Western societies?

Modern legal systems recognize different grounds for discrimination such as race, gender, religion, ethnicity, sexual orientation, age, and disability, based on the social and historical processes that shaped the constitutional traditions of each country. Since the second half of the twentieth century, international and regional human rights mechanisms have fostered a strong cosmopolitan culture against the exclusion of groups and individuals on the basis of physical and cultural attributes. Unfortunately, because discrimination is a structural feature of our societies, the path to equality is a never-ending challenge. In the context of this unfinished and ongoing process, several communities have raised a new equality claim: neurodiversity.

---

1. In a social marketplace dominated by communication and networking, by the recognition of a limited range of learning capacities, and by the celebration of status-oriented social conventions—all things rooted in brain functions and structures—that together with the environment contribute to people’s socioeconomic welfare—what happens to individuals who, by nature, struggle with these cognitive tasks?
Over the last two decades, a social movement has attempted to redefine the perception of brain-based disorders by reconsidering the nature of atypical perceptual and cognitive performance. This neurodiversity movement is an international civil rights effort that embraces a concept of cognitive traits and brain conditions using either a narrow or broad definition. More narrowly, some specialists and activists limit neurodiversity to individuals with High-Functioning Neurodevelopmental Disorders (NDDs) such as autism spectrum disorder (ASD). More broadly, others understand neurodiversity as referring to both "High" and "Low" Functioning autistics, as well as to a larger set of other disorders such as attention deficit/hyperactivity disorder, bipolar disorder, developmental dyspraxia, dyslexia, epilepsy, and Tourette Syndrome. On the whole, the neurodiversity movement encompasses different approaches. It challenges oppressive social norms, stigma, and rejection, and pleads for an expansion of support systems such as inclusion-focused services, accommodation, vocational training, and independent living support. Some of the movement's activists have more radical positions. They reject the idea that different neurodevelopmental profiles need to be cured, and consider NDDs a natural human variation rather than a pathology. This view radically shifts the focus of advocacy towards concepts such as human diversity and identity.


5. See the analysis of these neurodiversity claims in Baker, supra note 3, at 20–21; and Ortega, supra note 2, at 426.
In Western countries, sociopolitical issues are often framed in legal terms, and transformation is implemented through legal change.\(^6\) The efforts of the neurodiversity movement, that in general constitutes a branch of the disability right movement, do not diverge from these historical patterns, in which social actors relied upon the law and litigation. On the one hand, it pleads for a reconceptualization of the reductive labels traditionally applied to individuals with NDD; on the other hand, from the legal and judicial perspective, it demands an end to the exclusion of people with cognitive diversities and seeks equality. Thanks to this political and legal endeavor, the concept of neurodiversity may soon enrich the constitutional list of grounds of discrimination, and legal systems might embrace the idea of a brain-based equality or brain pluralism.

The autism rights movement is the most prominent component of the neurodiversity movement.\(^7\) In the last two decades, perceptions of the polymorphic manifestations of ASD have radically changed in several ways.\(^8\) First, people on the spectrum are breaking down the wall of silence surrounding their condition using technological communication advances. Second, experts are developing a more comprehensive scientific understanding of NDDs. Third, there has been significant growth in the number and strength of advocacy organizations, followed by an attendant increase in litigation and legal lobbying to establish and defend new rights.

In the face of these efforts, there are vocal critics of the neurodiversity movement. According to the harshest critics, re-conceptualizing brain conditions as a form of diversity, instead of as impairing disorders, would undermine the already insufficient services provided to this vulnerable population.\(^9\) This

---


\(^8\) Ian Hacking, Humans, Aliens & Autism, 138 DAEDALUS 44, 46–49 (2009).

\(^9\) See Jaarsma & Welin, supra note 2, at 27–28 (arguing that, by accepting neurodiversity as a special culture, the autists that need care will face a hard time receiving special treatment).
BRAIN EQUALITY

perspective embodies the contrast between the strict "medical model" of brain diversity, and the neurodiversity movement's rights and citizenship model. However, the neurodiversity movement does not underestimate the severely maladaptive consequences of these conditions. Rather, they argue that if society considers their diversity a disease, then society does not need to change—differences are subsumed into symptoms, and individuals become merely patients who must be cured in order to eradicate their abnormality whenever possible. If the medical model necessitates diagnosis and treatment to change the patient with divergent characteristics, the rights and citizenship approach, on the contrary, aims to modify the social environment to be more inclusive, and to provide better and more accessible physical and social re-habilitation. Their goal is enabling achievement of independence, inclusion, and participation in society by divergent people. The paradigm shift from a medical model towards a rights and citizenship model of disability does not mean that the health care system would be exempt from ensuring services. Instead, such a shift would reorient those services away from erasing diversity in order to achieve normality, towards helping people to achieve their fullest individual potential by alleviating social and physical difficulties they confront.

This article examines some of the implications of the concept of neurodiversity. First, it analyzes the definition of neurodiversity based on the existing literature and academic debates. Then, discusses the legal implications of neurodiversity. Third, it explores the claim made by a large part of the neurodiversity movement that people with different neurodevelopment profiles should be considered a new minority. Finally, this article discusses how neurodiversity might require a new interpretation of the idea of constitutional equality. In order to inquire into these issues, this article comparatively analyzes NDD-related jurisprudence of the highest North American and European courts over the last fifteen years. Examining these decisions helps determine the current relationship between neurodiverse individuals and legal sys-

10. See Baker, supra note 3, at 8–9 (describing the understanding of disability under a medical model and a social model); Jaarsma & Welin, supra note 2, at 24–25 (describing the medical model of interpreting behavior).
tems, highlighting that the principle of equality is under intense pressure when the context implicates neurological diversities.

II. NEURODIVERSITY: DEFINITIONS, AND DEBATES

The concept of neurodiversity refers to a highly variable cluster of atypical neurological mechanisms resulting in maladaptive behaviors.\textsuperscript{12} The term neurodiversity refers to a broad spectrum of cognitive, linguistic, and learning functions, and early-onset neurobiological conditions that impair the individual’s capacity for social understanding, social interaction, learning, and pragmatic and semantic communication. NDDs and pervasive developmental disorders (PDD) produce overreactions to environmental stimuli and, in most cases, repetitive or restricted interests and activities.\textsuperscript{13} Atypical neurological functioning is neither a unitary condition nor the result of a singular genetic, molecular or cellular etiology. Symptoms vary significantly in character and severity, occur in all ethnic and socioeconomic groups, and affect every age.\textsuperscript{14}

The term neurodiversity first appeared at the end of the twentieth century. It reached the wider public through an article by journalist Harvey Blume published in 1998, in which he stated that “[n]eurodiversity may be every bit as crucial for the human race as biodiversity is for life in general.”\textsuperscript{15} One year later, Judy Singer, an Australian sociologist, expanded on this neologism. She is generally credited with the term neurodiversity.\textsuperscript{16} Singer articulates a “new category of difference” in the

human population: individuals diagnosed with NDDs such as ASD and Asperger’s Syndrome. She suggests that “the neurologically different represent a new addition to the familiar political categories of class, gender and race.”

Over the last two decades, specialists in different fields analyzed this concept, proposing several definitions in the growing literature. According to Francisco Ortega: ‘‘neurodiversity’ asserts that some features usually described as illnesses are in fact only atypical or ‘neurodivergent’, i.e. they result from a specific ‘neurological wiring’. Therefore, it is merely a human difference that must be respected like any other such difference (be it sex, race or any other attribute).” Pieer Jaarsma and Stellan Welin point out:

What we call the neurodiversity claim consists of at least two parts. One is related to the idea that there are indeed neurological (or brain-wiring) differences among the human population. Being autistic is one of them. One aspect of the neurodiversity claim is that autism (or some other neurological condition) is a natural variation among humans. Being neurodiverse or neurotypical . . . are just different ways of existing as humans. The second aspect of the neurodiversity claim is related to rights, non-discrimi-

17. Judy Singer, ‘Why Can’t You Be Normal for Once in Your Life?’ From a ‘Problem with No Name’ to the Emergence of a New Category of Difference, in Disability Discourse 59 (Mairian Corker & Sally French eds., 1999).
19. Ortega, supra note 2, at 431.
nation and other more political issues. The two aspects often go together.\(^2\)\(^0\)

Christina Nicolaidis’ definition says:

The neurodiversity movement challenges us to rethink autism through the lens of human diversity. It asks us to value diversity in neurobiologic development as we would value diversity in gender, race, ethnicity, religion, or sexual orientation. As opposed to only focusing on impairments, the neurodiversity model sees autistic individuals as possessing a complex combination of cognitive strengths and challenges.\(^2\)\(^1\)

Whereas Chon-Ming Lim affirms:

For the past decade, activists and theorists . . . have been trying to reconceptualize autism. One of their central claims is that autism is not a disorder. For them, autistic traits are the result of atypical (rather than abnormal) neurological structures, which give rise to different types and levels of functioning from those arising from the structures in neurotypical individuals.\(^2\)\(^2\)

For the purpose of this article, Dana Lee Baker’s definition depicts this concept in the most precise way:

[N]eurodiversity refers to atypical functionalities found in individuals who have identifiable neurobiologic differences and to their interactions with individuals considered neurologically typical in the context of public infrastructures built around a presumption of neurotypicality. For the most part, this implies that communities referred to as neurodiverse include only those that incorporate individuals who have been formally diagnosed (or could be, given access to professionals) with a disability believed to involve a significant brain-based difference compared to what is currently considered the human norm.\(^2\)\(^3\)

\(^2\)\(^0\) Jaarsma & Welin, supra note 2.
\(^2\)\(^1\) Christina Nicolaidis, What Can Physicians Learn from the Neurodiversity Movement?, 14 Virtual Mentor 503 (2012).
\(^2\)\(^2\) Lim, supra note 3, at 564–65.
\(^2\)\(^3\) Baker, supra note 3, at 19.
The concept of neurodiversity, however, generates disagreement. According to neurodiversity critics, severe disruption in otherwise typical human functions cannot be defined as anything other than pathology. For them, as promising as the neurodiversity discourse might be, severely maladaptive ASD phenotypes are medical conditions. They maintain that nosology, the branch of medical science that classifies and describes the characteristics of diseases, is not a social construction. They assert that symptoms deeply impact people's lives, and disabling functionalities are ill-defined as mere human diversity. Moreover, according to neurodiversity skeptics, the shift away from the medical model and the contextual emphasis on diversity might trigger dangerous outcomes. Francisco Ortega reviewing the complexity of the concept of neurodiversity, states: “parents [are] fighting to obtain governmental support or make health insurance companies pay for the therapy. . . . The advocates’ claim that autism is not an illness and that attempts to cure it violate autistic rights may therefore provide a legal backing for refusing to subsidize the therapy.” Furthermore, critics of the neurodiversity movement insist that opposition to the idea of the neurotypical world would not help to increase social awareness of NDDs’ implications. According to skeptics, the diversity discourse should mobilize only in reference to individuals who display high-functionality or to NDDs manifesting less intrusive profiles, such as language-based learning disabilities or attention deficits. Although identity-building agendas are understandable reactions to against past marginalization, skeptics argue that high-functioning individuals might monopolize the

24. See id. at 27–44 (describing tensions and disagreements between four competing neurodiversity-related conceptions and policy agendas: cause, care, cure, and celebration).

25. This polarization is analyzed in Nancy Bagatell, From Cure to Community: Transforming Notions of Autism, 38 ETHOS 33 (2010); Steven K. Kapp et al., Deficit, Difference, or Both? Autism and Neurodiversity, 49 DEVELOPMENTAL PSYCHOLOGY 59 (2013); and Chloe Silverman, Fieldwork on Another Planet: Social Science Perspectives on the Autism Spectrum, 3 BIOSOCIETIES 325 (2008).

26. Ortega, supra note 2, at 429.


neurodiversity discourse at the expense of the most vulnerable, who would then see NDD health-care dangerously downsized.\textsuperscript{29}

Today, the dichotomy of diversity versus pathology appears less rigid, as debates show a variety of nuanced positions.\textsuperscript{30} Advocates who emphasize the idea of diversity do not deny the importance of health care services, including educational and behavioral individualized interventions.\textsuperscript{31} Instead, they argue that by acting immediately on parental concerns, monitoring behavior and development, asking promptly for a comprehensive evaluation, searching for etiologic and co-morbid conditions, and managing medical issues, neurodiverse people increase their chances of living independently as adults even without being cured.\textsuperscript{32}

\section*{III. Legal Neurodiversity}

Despite the ongoing social debates surrounding the subject, the concept of neurodiversity displays substantial legal attributes. Advocates' perspective of diversity proposes expansion of protections for neurodiverse individuals by triggering anti-discrimination mechanisms already in place for other human diversities. Neurodiversity also renews debates around the principle of constitutional equality, and fosters a reinterpretation of the theory of justice.\textsuperscript{33}

Importantly, neurodiversity from a legal perspective does not inherently conflict with the medial approach and its focus

\textsuperscript{29} See, e.g., Jaarsma & Welin, supra note 2 (noting that acceptance of neurodiversity as a separate culture may not be desirable for low-functioning autists who need care).

\textsuperscript{30} Evrinomy Avdi, Negotiating a Pathological Identity in the Clinical Dialogue: Discourse Analysis of a Family Therapy, 78 PSYCHOL. & PSYCHOTHERAPY 493 (2005); Fenton & Krahn, supra note 4, at 4; Ortega, supra note 2, at 441.

\textsuperscript{31} Editorial, Pride in Autistic Diversity, 387 LANGET 2479 (2016).

\textsuperscript{32} Caruso, supra note 7, at 514–37; Jaarsma & Welin, supra note 2; Kapp et al., supra note 25, at 68; Scott Michael Robertson, Neurodiversity, Quality of Life, and Autistic Adults: Shifting Research and Professional Focuses onto Real-Life Challenges, 30 DISABILITY STUD. Q. no. 1, 2009, http://dx.doi.org/10.18061/dsq.v30i1.1069; Editorial, supra note 31, at 2477.

\textsuperscript{33} Martha Nussbaum, The Capabilities of People with Cognitive Disabilities, 40 METAPHILOSOPHY 331 (2009). See generally MARTHA C. NUSSBAUM, FRONTIERS OF JUSTICE: DISABILITY, NATIONALITY, SPECIES MEMBERSHIP 1 (2006) (identifying “people with physical and mental impairments” as one of the “three unsolved problems of social justice”).
on the pathological dimensions of NDDs. In the legal realm, the two dimensions complement each other: neurodiverse individuals can be considered patients and unfairly excluded citizens. As such they may receive more strategies to implement in the public sphere. While people with severely disrupted functioning need more efficient medical systems and might prefer to have symptoms medicalized, less affected individuals might prefer to focus on the identity-diversity level and demand a more inclusive society. Legal neurodiversity respects this individual choice. By adding the discrimination-equality dimension, the concept of legal neurodiversity aims to create more inclusive communities by fostering social change and accommodations. In short, considering the legal implications of neurodiversity suggests that the idea of better-assisted patients is not incompatible with the one of more empowered citizens.

The legal dimension of neurodiversity has theoretical and pragmatic features. From a theoretical perspective, overcoming stigmatization and labeling is the pre-condition for avoiding discrimination. From a pragmatic perspective, three variables arise for consideration: 1) the number of neurodiverse people living in our societies—prevalence data suggest that atypical neurological conditions are sufficiently common in our societies to merit attention; 2) the extent to which people impacted by NDDs struggle to comply with or fit into legal norms and behavioral standards, as well as the legal consequences of this gap; and 3) the growing number of neurodiversity-related litigation cases. This article considers each individually below.

A. The Size of the Neurodiverse Population

Considering just ASD alone, the U.S. Center for Disease Control and Prevention (CDC) released data finding that one in 59 children has an ASD. The U.S. National Health Inter-

view Survey, a nationwide population study based on in-person, household, and parent interviews, determined that the 2014–2016 ASD prevalence was 2.41%. A 2011 study in Stockholm found similar frequency: they found prevalence of ASD of 0.40%, 1.74%, 2.46%, and 1.76% among 0–5, 6–12, 13–17 and 18–27 year olds respectively. The corresponding proportion of individuals with an intellectual disability was 17.4%, 22.1%, 26.1%, and 29.4%. Between 2001 and 2011, ASD prevalence increased almost 3.5 fold among children aged 2–17 years. To explain the increase, researchers highlight an eightfold increase of ASD without intellectual disability, from 0.14 to 1.10 %, while the prevalence of ASD with intellectual disability increased from 0.28 to 0.34%. Authors affirm that the soar in ASD prevalence “is likely contributed to by extrinsic factors such as increased awareness and diagnostics.”

If we include other types of atypical neurological functioning in the concept of neurodiversity, prevalence increases dramatically. According to the International Dyslexia Association, between 15% and 20% of the population displays language-based learning disabilities such as dyslexia, dysgraphia,
and dyscalculia. These learning disabilities affect males and females nearly equally, and rates are consistent across different ethnic and socio-economic backgrounds. In the United States, studies suggest that between 5% and 10% of the population display dyslexic traits alone. In languages with orthography and phonology exceptions, like Danish and English, dyslexic traits reach up to 12% of the population. With languages like Italian, German, and Dutch, in which there is a higher grapheme-to-phoneme correspondence, prevalence in the population is lower but consistent. According to a U.S. report published in 2014 by the Center for Learning Disability, this population typically suffers from low self-esteem, sets low expectations for themselves, struggles with underachievement and underemployment, and engages with the criminal justice system with greater frequency than their non-learning disorder peers.

Although severely maladaptive, NDDs without intellectual impairments, and other atypical neurological conditions such as learning disorders (LDs) that are not associated with diminished intelligence are also widely prevalent. ASD associated with intellectual disability, for example, only comprises about 30% of the cases of diagnosed autism today, and impair-
ments are consistently determined by co-morbid factors. A study by the Californian Department of Developmental Services (DDS) confirms this figure. According to its 2007 service survey, only 36% of the individuals assessed on the ASD spectrum have mental retardation diagnoses. This percentage encompasses different levels of mental disability, varying from severe to mild. Although this finding contradicts previous prevalence rates, the California data parallels the most recent literature.

Based on the above-mentioned prevalence data, both in the United States and in Europe, several millions of adolescents displaying a wide range of neurodiverse conditions, ranging from NDDs to LDs, will transition to adulthood over the next decade. This will significantly impact the labor market as well as health care and social security systems. Atypical neurological conditions are a structural feature of the human population. The earliest identified cases of autism were of this type; however, the later identified autism without intellectual impairment is notably distinct, with different genetic causes and much lower levels neurological damage. See Michael Ronemus et al., The Role of De Novo Mutations in the Genetics of Autism Spectrum Disorders, 15 Nature Revs.: Genetics 133, 133 (2014) (discussing de novo mutations as a cause of ASD).


48. Id.


50. See Elise B. Barbeau et al., The Level and Nature of Autistic Intelligence III: Inspection Time, 122 J. Abnormal Psychol. 295, 295 (2013) (discussing how different assessment methodologies lead to different results). Inspection time (IT) is a processing speed measure associated with general intelligence in typical individuals. Compared with a Wechsler IQ-matched typical group, the autistic group’s visual IT was significantly shorter, a processing speed advantage that vanished when Raven’s Progressive Matrices (RPM) was used for group matching. Id.; see also Idring et al., supra note 36, at 4 (discussing co-morbid rates of autism and intellectual disabilities).

ulation. Atypical brain structure and functioning are quantitatively and qualitatively relevant in our societies, crafting a demographic and anthropological reality that cannot be ignored.\textsuperscript{52}

B. Legal Implications of Neurodiversity

The question of what the legal consequences of these widespread human cognitive discrepancies are remains.\textsuperscript{53} It is unclear whether legal systems are equipped to deal with this previously underestimated cognitive pluralism. Diverse cogni-

\textsuperscript{52} A deepening understanding of the brain is fostering a re-conceptualization of behavioral anthropology as well as of the range and roots of, and the proper response to, an increasing awareness of neurodiversity. Owen D. Jones et al., Neuroscientists in Court, 14 Nature Revs.: Neuroscience 730 (2013) (discussing the increasingly prevalent role of neuroscientific evidence in courts). New ideas are transforming social science heuristic aspects, vocabulary, and research methodologies. See, e.g., Neurocultures: Glimpses into an Expanding Universe (Francisco Ortega \& Fernando Vidal eds., 2001) (discussing the cerebral subject); The New Brain Sciences: Perils and Prospects (Dai Rees \& Steven Rose eds., 2004) (discussing neurobiologization of selfhood); Nikolas Rose, The Politics of Life Itself: Biomedicine, Power, and Subjectivity in the Twenty-First Century (2007) (outlining the neurochemical self); Alain Ehrenberg, Le Sujet Cérébral, 309 Esprit 130 (2004) (cerebral subjectivation); see also Ortega, supra note 2, at 426 (discussing neurodiversity and identity). This movement has developed into a new field of interdisciplinary research: social neurosciences. See, e.g., Louis Cozolino, The Neuroscience of Human Relationships: Attachment and the Developing Social Brain (2d ed. 2014) (exploring the cognitive effects of interpersonal relations); Ralph Adolphs, Cognitive Neuroscience of Human Social Behavior, 4 Nature Revs.: Neuroscience 165 (2003) (exploring the cognitive underpinnings of interpersonal relations); Editorial, Focus on Social Neuroscience, 15 Nature Neuroscience 645 (2012) (discussing then-current progress in the field of social neuroscience); Matthew D. Lieberman, Social Cognitive Neuroscience: A Review of Core Processes, 58 Ann. Rev. Psychol. 259 (2007) (offering a broad overview of key research domains and processing distinctions in social cognitive neuroscience). Over the last decade, studies conducted by scientists and legal scholars have opened a new field of interdisciplinary research probing the intersection of neurosciences and the law. Among a growing literature, see The Impact of Behavioral Sciences on Criminal Law (Nita A. Farahany ed., 2009); International Neurolaw: A Comparative Analysis (Tade Matthias Spranger ed., 2012); Law and the Brain (Semir Zeki \& Oliver Goodenough eds., 2006); Law, Mind and Brain (Michael Freeman \& Oliver R. Goodenough eds., 2009); Michael S. Gazzaniga, The Law and Neuroscience, 60 Neuron 412 (2008); Michael S. Pardo \& Denis Patterson, Minds, Brains, and Norms, 4 Neuroethics 179 (2011).

tion, diverse sociality, diverse learning, and diverse stimuli processing exhibited by a large population may deeply affect expected behaviors or intercommunication conventions based on commonly accepted legal and social norms. This may alter the perception and the understanding of behaviors displayed by neurodiverse individuals with dangerous consequences such as: "mannerisms exhibited by a person suffering from [ASD similar disorders] are very similar to reactions associated with a guilty mind" as well as restricted or obsessive-compulsive interests may be considered criminal offence or triggering unmotivated suspicions. Moreover, difficulty with sensory overload, semantics, sarcasm, changes in routine or structure, poor social awareness, and inadequate understanding of nonverbal communications such as body language and facial expressions, are all traits that may create severe responses within society and attendant costs for individuals with


56. Neurosciences are providing evidence that brain structure and functioning, which determine perceptual and cognitive performance, emotions, social and moral thought, and decision making, are both governed by universal principles and deviations. Discrepancies are visible both in typical and atypical population in relation to age, sex, personality, culture and genetics. At the same time, neurosciences are providing the "first compelling evidence that neuropsychiatric and neurodevelopmental disorders reflect fundamental differences in brain structure and function." John D. Gabrieli et al., Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience, 85 NEURON 11, 11 (2015). Discrepancies affect the capacity to learn, to judge and to respond to behavioral and pharmacological treatments. They impinge upon the capacity to solve cognitive conflicts implicated in impulsivity (behavioral disinhibiting) and to make decisions. Id. While diversities are genetically based or determined by the onset of pathologies, socio-cultural interactions are fundamental in shaping the scale of brain diversity one may carry. Id. at 21-22.


NDDs.\textsuperscript{59} Since the description of autism’s phenotype in the first half of the twentieth century, the understanding of NDDs has advanced and its legal implications are better understood.\textsuperscript{60} However, the interaction between those conditions and the law still remain problematic.\textsuperscript{61}

Studies reveal a dramatic relationship between NDDs and the criminal justice system. In 2001 the National Research Council released a report demonstrating that “individuals with neurodevelopmental disabilities are involved, in one role or another, in the criminal justice system at a much higher rate than persons without disabilities.”\textsuperscript{62} The findings of a study conducted in the UK “confirm the presence of a significant number of people with NDD in a male prison” where 87 inmates out of 240 in a prison in London were affected.\textsuperscript{63} A study conducted in Sweden found 13% of inmates in Swedish forensic settings have ASD.\textsuperscript{64} In North America, a study estimated the prevalence of ASD in U.S. prison to be 4.4%, four


\textsuperscript{60} Caruso, supra note 7, at 483; Cohen et al., supra note 51, at 419.


\textsuperscript{63} J. McCarthy et al., Characteristics of Prisoners with Neurodevelopmental Disorders and Difficulties, 60 J. Intell. Disability Res. 201, 201 (2015).

times greater than in the general population—out of 1800 inmates in a maximum-security state prison in the Midwest.\(^{65}\) In 2015, the U.S. National Council of Disability released a report concluding that many disabled youth in the U.S. juvenile justice and criminal justice systems are deprived of an appropriate education that could disrupt the so-called "School-to-Prison Pipeline." Among incarcerated youth, 85% have learning and or emotional disabilities, yet only 37% receive special education.\(^{66}\) Most of these youth are either undiagnosed or not properly served in school.\(^{67}\) Many students have invisible disabilities, such as specific LDs,\(^{68}\) emotional disturbance, posttraumatic stress disorder, or attention deficit/hyperactivity disorder.\(^{69}\) Moreover, a 2004 survey on state and federal inmates in the United States revealed that 31% self-reported having speech and reading difficulties and 43% reported taking special education classes.\(^{70}\) A study conducted in Finland in 2014 determined that between 29% and 36% of the inmates displayed reading and spelling disorders.\(^{71}\) In the UK, a screening performed in a British institution in 2004 found high levels of speech, language, and communication difficul-


\(^{66}\) Nat'l Council on Disability, Breaking the School-to-Prison Pipeline for Students with Disabilities 6 (2015).

\(^{67}\) Id.


\(^{69}\) Nat'l Council on Disability, *supra* note 66, at 6.

\(^{70}\) Jennifer M. Reingle Gonzalez et al., *Disproportionate Prevalence Rate of Prisoners with Disabilities: Evidence from a Nationally Representative Sample*, 27 J. Disability Pol'y Stud. 106, 106, 112 (2016) ("Disability prevalence remained substantially higher among prisoners than among the non-institutionalized population. Prisoners were more likely to report specific learning, sensory, and speech-related disabilities than non-institutionalized adults.... In summary, 41% of prisoners reported a disability, most commonly, learning disabilities. Prisoners with disabilities were identified as an at-risk group for recidivism, given their pre-incarceration experiences, and limited vocational and work-related training received in prison.").

ties among young offenders. Multiple studies in northern Europe show an overrepresentation of reading and writing disabilities among inmates in which the prevalence varies between 6 and 70%. The reason for this problematic discrepancy is the difficulty in defining and diagnosing reading-writing disabilities and dyslexia. More research on prevalence and interventions is greatly needed.

Although fragmented and incomplete, the available data nevertheless shows a disproportionately high number of neurodiverse people in prison populations. As several law and criminology scholars point out, this is not determined by intrinsic higher criminal dispositions, but rather by the problematic interaction between NDDs and the criminal justice. A study published in 2007 found "little evidence to support the notion that offending was a significant problem in people with


73. Idor Svensson, Reading and Writing Disabilities Among Inmates in Correctional Settings: A Swedish Perspective, 21 LEARNING & INDIVIDUAL DIFFERENCES 19, 23 (2011). See also Grigorenko, supra note 68, at 355–56 (reviewing studies conducted in the United States and in Sweden that found an elevated frequency of learning problems among juvenile delinquents); May Lindgren et al., Dyslexia and AD/HD Among Swedish Prison Inmates, 3 LEARNING & INDIVIDUAL DIFFERENCES 19, 23 (2011). See also Grigorenko, supra note 68, at 355–56 (reviewing studies conducted in the United States and in Sweden that found an elevated frequency of learning problems among juvenile delinquents); May Lindgren et al., Dyslexia and AD/HD Among Swedish Prison Inmates, 3 LEARNING & INDIVIDUAL DIFFERENCES 19, 23 (2011).

74. Claire King & Glynis H. Murphy, A Systematic Review of People with Autism Spectrum Disorder and the Criminal Justice System, 44 J. AUTISM & DEVELOPMENTAL DISORDERS 2717, 2721 (2014); Matthew D. Lerner et al., Emerging Perspectives on Adolescents and Young Adults with High-Functioning Autism Spectrum Disorders, Violence, and Criminal Law, 40 J. AM. ACAD. PSYCHIATRY & L. 177, 178 (2012); see also Ann Browning & Laura Caulfield, The Prevalence and Treatment of People with Asperger's Syndrome in the Criminal Justice System, 11 CRIMINOLOGY & CRIM. JUST. 165 (2011) (surveying the relation between ASD and criminality); Catherine A. Cheely et al., The Prevalence of Youth with Autism Spectrum Disorders in the Criminal Justice System, 42 J. AUTISM & DEVELOPMENTAL DISORDERS 1856 (2011) (comparing rates of criminality between youths with and without ASD).

75. See Sissel Berge Helverschou et al., Personal Experiences of the Criminal Justice System by Individuals with Autism Spectrum Disorders, 22 AUTISM 460, 466 (2018) (documenting the experiences of offenders with ASD in the criminal justice system).
Asperger.” Although different opinions suggest that individuals with this condition appear more frequently in forensic populations than the general public, other studies affirm that “individuals with autism are easily manipulated, and therefore easily enticed into criminal behavior” and that co-morbid psychiatric factors might determine unconventional behaviors. As Cohen, Dickerson and Forbes point out:

We have a general understanding today that even individuals with high-functioning ASDs can have issues with sensory overload, semantics, sarcasm, have difficulty when dealing with changes in routine or structure, generally have poor social awareness, and inadequate understanding of nonverbal communications such as body language and facial expressions, both on the giving and receiving end. Accordingly, their responses within society can be very difficult and they can often be viewed as exhibiting antisocial behavior. While there is no evidence to suggest that individuals with ASDs will commit crimes at a higher rate than the general population, it stands to reason that offenses may be committed by individuals with ASD. However, as with any offense, intent must be considered, and when considering an individual with an ASD, such intent may have to be evaluated differently.

80. Cohen et al., supra note 51, at 413 (citations omitted).
Thomas Mayers presents an even more radical point of view on this issue, suggesting, "[t]o the extent that persons with autism are ill-prepared to conform to society's demands, their newfound membership in society is a cruel illusion: the substitution of incarceration for institutionalization."81

Two general aspects frame the debate on the legal neurodiversity. First is the relationship between constitutional principles and neurodiversity, especially equality and non-discrimination. Courts of different countries deal with an increasing number of cases involving autism. While it is difficult to determine accurately the number of cases in Europe due to the diversity of national legal systems, the United States provides a good example of the phenomenon: between 2000 and 2017, U.S. federal courts decided more than 2700 cases in which the word “autism” appeared.82 During the same two decades, more than 700 cases mentioning the words “Asperger Syndrome” were decided.83 Within these U.S. cases, a growing number of plaintiffs claimed violations of the equal protection clause of the Fourteenth Amendment of the U.S. Constitution and of the due process clause of the Fifth Amendment. The equality principle was also repeatedly challenged in European jurisdictions.84

The second aspect is the impact of the UN Convention on the Right of Persons with Disabilities (CRPD) and its Optional Protocol adopted in New York on December 13, 2006.85 This Convention is the first international treaty that defines the specific rights of people with disabilities, and thus constitutes a paradigm shift in disability rights.86 While the Convention

81. Mayes, supra note 61, at 98.
84. See DEBORAH MABBETT, BRUNEL UNIV., DEFINITIONS OF DISABILITY IN EUROPE: A COMPARATIVE ANALYSIS 67 (2002) (discussing discrimination and equality in Ireland, the United Kingdom, Sweden, Norway and France).
86. See Michael Ashley Stein et al., Health Care and the UN Disability Rights Convention, 374 LANCET 1796, 1796 (2009) (noting that the CRPD is the first
does not list any specific disabilities, given the evolving nature of diagnostic labels and concepts of disability, this document specifies that "[p]ersons with disabilities include those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society in an equal basis with others."\(^8\)\(^7\) Article 3 establishes general principles that are now vital in dealing with neurodiversity:

1) Respect for inherent dignity, individual autonomy including the freedom to make one's own choices, and independence of persons; 2) Non-discrimination; 3) Full and effective participation and inclusion in society; 4) Respect for difference and acceptance of persons with disabilities as part of human diversity and humanity; 5) Equality of opportunity; 6) Accessibility; 7) Equality between men and women; 8) Respect for the evolving capacities of children with disabilities and respect for the right of children with disabilities to preserve their identities.\(^8\)\(^8\)

The CRPD also establishes an implementation monitoring mechanism through the Committee on the Right of Persons with Disabilities.\(^8\)\(^9\) All European states have signed and ratified the CRPD and the Optional Protocol, and the European Union and the Council of Europe are now official members of this international mechanism.\(^9\)\(^0\) In contrast, the United States

---

\(^8\)\(^7\) G.A. Res. 61/106, \textit{supra} note 85, art. 1.

\(^8\)\(^8\) \textit{Id.} art. 3.

\(^8\)\(^9\) \textit{Id.} arts. 33–34 (mandating that state parties implement monitoring programs and other governance activities for the rights of persons with disabilities).

\(^9\)\(^0\) \textit{See} Jarlath Clifford, \textit{The UN Disability Convention and its Impact on European Equality Law}, 6 \textit{EQUAL RTS. REV.} 11, 11 (2011) (noting that this is the first time in history the European Union has ratified an international human rights treaty that has subsequently influenced the European Union to a significant degree).
is one of the few countries that signed Convention, but has not yet ratified it.\textsuperscript{91}

IV. NEURODIVERSE INDIVIDUALS AS A NEW MINORITY?

In line with Hans Asperger's 1938 statement that autistic traits are "not rare at all," Steven Silberman claims that "given current estimates of prevalence, autistic people constitute one of the largest minorities in the world."\textsuperscript{92} The Autistic Self Advocacy Network (ASAN) affirms that: "[l]ike any other minority group, we have the right to respectful and equal treatment in all aspects of society,"\textsuperscript{93} and "[a]s with other minority groups, an inclusive educational system is a vital part of changing social attitudes and creating a culture in which all people are seen as equal participants in society."\textsuperscript{94} Other similar statements coming from within the neurodiverse community consider autistic people a distinct minority group.\textsuperscript{95}

While an official legal definition of minority groups does not exist internationally, specialists and UN documents agree that identifying a group as a minority depends on both objective, such as the existence of a shared ethnicity, language or religion, and subjective—individuals must identify themselves as members of a minority—criteria being met.\textsuperscript{96} With this, international law and state constitutions recognize the existence of minority groups living within their territory based on classic objective criteria.\textsuperscript{97} However, over the last decades, the prob-


\textsuperscript{92.} SILBERMAN, supra note 18, at 469.


\textsuperscript{94.} Id.

\textsuperscript{95.} See, e.g., Amy Nelson, Declaration from the Autism Community that they see [sic] a Minority Group, PRWEB (Nov. 18, 2004), http://www.prweb.com/pdfdownload/179444.pdf; see also Jaarsma & Welin, supra note 2.


\textsuperscript{97.} In 1977, Francesco Capotorti, as UN Special Rapporteur, defined a minority as: "[a] group numerically inferior to the rest of the population of a State, in a non-dominant position, whose members—being nationals of the
Problem of defining new minority groups—such as gender, LGBT, and disabled people—came to the forefront, and literature established a set of guidelines based on additional objective and subjective criteria for identifying new minorities needing protection against discrimination and redress for past injustices. According to these guidelines, in order to identify as a new minority, a group of people needs: 1) to have significantly less control or power over their lives than members of a dominant or majority group; 2) not to be limited to a mathematical minority; 3) to be interchangeable with a subordinate group; 4) to experience a narrowing of opportunities, such as access to education, jobs, or health care system, compared to society; 5) to suffer discrimination and subordination; 6) to manifest physical and or cultural traits that set them apart, and which are disapproved by the dominant group; 7) to share a sense of collective identity and common burdens, 8) to share social rules around membership; 9) to tend to marry within the group. According to these criteria, states may not only protect linguistic, ethnic, or religious minorities within their territories and jurisdictions, but also new collective aggregations of citizens sharing visible common attributes who have a common history of unfair treatment.

Under these guidelines, neurodiverse individuals have good reasons for perceiving themselves as a minority and seeking official state recognition as such. According to guidelines 1, 2, 3, and 6 above, NDDs individuals constitute a subordinate group with common traits differing from the general population. However, most importantly, individuals with NDDs are definable as a new minority because of their vulnerable status as outlined in guideline number 4 above. To that end, they

State—possess ethnic, religious or linguistic characteristics differing from those of the rest of the population and show, if only implicitly, a sense of solidarity, directed towards preserving their culture, traditions, religion or language.” Francesco Capotorti (Special Rapporteur of the Sub-Comm’n on Prevention of Discrimination and Protection of Minorities), Study on the Rights of Persons Belonging to Ethnic, Religious and Linguistic Minorities, ¶ 568, U.N. Doc. E/CN.4/Sub.2/384/Rev.1 (1979).

suffer from higher rates of victimization. Poor school achievement, "poor social understanding or circumscribed interests; difficulties in adjusting to the diagnosis; and the impact of social exclusion," all contribute to an increased risk of NDD individuals becoming victims or perpetrators of crimes. Furthermore, in accordance with guideline number 4 above, NDDs individuals belong to a group that experiences a disadvantaged socio-economic position. In the United States, health care data surveys reveal that people with intellectual and developmental disabilities meet most of the governmental criteria to qualify as a "medically underserved populations" (MUP). Moreover, out of 2000 autistic adults surveyed in 2016 by the National Autistic Society in the UK, only 16% have full-time paid work and only 32% have some kind of paid work. Full and part-time combined, compared to 47% of disabled people and 80% of non-disabled people. Over three quarters, 77%, of unemployed adults with ASD say they want

99. A study found that children with ASDs are bullied nearly five times as often as their peers, with approximately 46% of ASD children in middle and high school reporting to their parents that they had been victimized at school. Paul R. Sterzing et al., Bullying Involvement and Autism Spectrum Disorders, 166 Archives Pediatrics & Adolescent Med. 1058, 1058 (2012); see also Benjamin Zablotsky et al., Involvement in Bullying Among Children with Autism Spectrum Disorders: Parents’ Perspectives on the Influence of School Factors, 37 Behav. Disorders 179, (2012) at 180 (identifying children with Asperger syndrome as having heightened risk of involvement in bullying).

100. It is consistently clear that individuals with NDDs have difficulty picking up social cues and understanding other individuals’ thoughts and intentions, and they are vulnerable to a range of crimes from fraud, theft, and domestic violence to more violent crimes. See Christine N. Cea, supra note 57.

to work, and four in ten say they have never worked. All of these factors clearly confirm the serious fragility and unprivileged conditions of this population, suggesting the neurodiverse population is vulnerable and deserves to be considered a new minority group.

Although the intrinsic concept of minority entails a numerically smaller entity compared to a majority, it does not mean a demographically negligible size. As a relatively substantial group of people with shared qualities, minorities are more than visible. Neurodiversity prevalence data show that minorities are a numerically consistent population living in our societies. Therefore, the claim that neurodiverse people constitute a new minority does not discredit the fact that they are not a small number of individuals affected by rare symptoms. On the contrary, the minority-group discourse embraced by the neurodiversity community aims to redress the idea of the supposed low-prevalence of NDDs.

However, minority status creates two substantial challenges. First is the issue of determining whether an individual belongs in the group. There are two strategies available for making this determination. Medical diagnosis and assessments may be used, but this approach could have dangerous repercussions. A diagnosis would then acquire additional symbolic meaning as the assessment of specific neuro-biological traits would have consequences beyond the realm of health care. Moreover, using purely diagnostic mechanisms for defining people as part of this specific minority would generate the effect of minority by force, in which external factors determine a person’s inclusion, regardless of individual will or identity.

Alternatively, society could rely on individual self-declaration.


103. Dark pages of human history have been written on the use of this mechanism. See, e.g., PIERRE-ANDRÉ TAGUIEFF, LA FORCE DU PRÉJUGE (1990) (discussing the near impossibility of individuals overcoming socially determined categorizations, namely, race); see also ETIENNE BALIBAR & IMMANUEL WALLERSTEIN, RACE, NATION, CLASS: AMBIGUOUS IDENTITIES 71–73 (1991) (discussing the classification of individuals in South Africa according to the legal categorizations prescribed by the government during Apartheid).
of belonging, which is a more flexible scenario emphasizing the intimate free-will aspect of this decision. The second challenge lies in understanding how much individuals with divergent profiles struggle in living their lives. This is important in order to determine the extent of protection required, but inevitably will vary in a NDDs population.104

The minority framework has at least two results. First, greater attention to minority experience demonstrates that legal norms, social conventions, and *habitus* are often established on the false assumption that they reflect overwhelmingly dominant human attributes,105 while minority rights efforts call for a redefinition of the concept of normalcy.106 Second, recognition of neurodiversity as a minority status necessary imposes a political dimension on NDD prevalence. Although people with different neurodevelopmental profiles are demographically relevant, lawmakers and politicians often ignore their needs and characteristics. The famous slogan "nothing about us without us" used by the ASAN demonstrates the political implications of this discourse.107 By asking for formal recognition as a new minority, neurodiverse people ask for participation, representation, and inclusion.108 This is a natural conceptual complement of the right and citizenship model of disability as opposed to a rigid medical one. The former implies active participation through empowerment, whereas the latter entails passivity.

104. See Designation of Medically Underserved Populations and Health Professions Shortage Areas, 75 Fed. Reg. 26,168, 26,168 (proposing reforms of the criteria for the Medically Underserved Population (MUP) designation); Kornblau, supra note 101, at 9 (arguing for the inclusion of people with intellectual and developmental disabilities within its MUP definition to increase access to healthcare and reduce poorer health outcomes); Richards, supra note 102 (proposing potential solutions to remediate exclusion of employees with Asperger syndrome).


106. See Baker, supra note 3, at 111 ("No tension is more central to modern disability policy development than the choice to pursue or redefine normalcy . . . . [N]egotiating a balance between desirable and non-desirable aspects of functional difference remains a painfully unsettled and unsettling topic in public and political discourse on all disabilities.").

107. See ASAN’s slogan in http://autisticadvocacy.org/

108. See supra Section 4 (Neurodiverse Individuals as a New Minority?).
V. HIGHEST COURTS AND NEURODIVERSITY-RELATED LITIGATION IN A COMPARATIVE PERSPECTIVE

In investigating the legal implications of neurodiversity, engaging in a comparative analysis of high court decisions from different jurisdictions that deal with neurodiversity is helpful. As discussed above, neurodiversity encompasses many neurological profiles, each corresponding to a large number of psychiatric diagnoses. Employing a broad definition of neurodiversity from a legal and judicial perspective would complicate and implicate the vast realm of mental health-related jurisprudence. From a judicial perspective it is hard to draw a clear distinction between mental disability and neurodiversity. This article takes a pragmatic approach: because the concept of neurodiversity is rooted in the autism rights movement, this article focuses on highest court jurisprudence on NDDs and ASD-related cases in order to verify the emergence among the most influential jurisdictions of neurodiversity-focused litigation, in a narrow sense.

The United States Supreme Court directly dealt with neurodiversity-related litigation on three occasions: *Winkelman v. Parma City Sch. Dist.*,109 *Bruesewitz v. Wyeth*,110 and *Endrew F. v. Douglas County School District RE-1*.111 The Canadian Supreme Court, in *Auton v. British Columbia*,112 decided an important case on autism, and while the Canadian Human Rights Tribunal engaged with the issue in *Dawson v. Canada Post Corporation*.113 The United Kingdom Supreme Court has decided at least two major cases involving autism, most notably *P v. Cheshire & Chester Council*,114 and *A v. Essex County Council*.115 The Spanish and the German Constitutional Tribunals decided two cases in 2014 involving equality and the right of education for ASD children.116 The European Court on Human Rights decided two cases in 2014 involving equality and the right of education for ASD children.

---

Rights, in the context of extensive jurisprudence concerning mental health issues, directly dealt with neurodiversity related implications in at least three major cases: Storck v. Germany,\(^\text{117}\) H.L. v. United Kingdom,\(^\text{118}\) Nelissen v. Netherlands,\(^\text{119}\) whereas the European Committee of Social Rights decided International Association Autism-Europe (IAAE) v. France.\(^\text{120}\) Although these courts embody diverse legal systems with different powers and jurisdictions, all of them have strong intra-systemic authoritative effects. Drawing on all of these cases, this article highlights the impact of neurodiversity-related case law in three main areas of the law: 1) education, health care, and social services; 2) physical liberty, competency, and criminal law; and 3) discrimination at the workplace.

A. Education, Health Care, and Social Services

Neurodiversity-related cases are often at the crossroads of health care and educational issues as behavioral interventions for diverse neurological conditions involve health and educational therapies. Consequently, this is one of the most litigated areas related to neurodiversity.

Although characterized by contradictions and systemic inequalities, the United States has seen a dramatic boost in the debate around neurodiversity like no other country in the world, mostly driven by the Autistic Self Advocacy Network and similar organizations.\(^\text{121}\) Under the pressure of public opinion, Congress has passed a number of new laws in this field over the last two decades.

The Children Health Act of 2000\(^\text{122}\) was the first tangible victory of the autism movement at the federal level. Under this statute, federal agencies must undertake a long-term study of

\[^{121}\text{Caruso, supra note 7, at 487; Sheryl Dicker & Emily Bennett, Engulfed by the Spectrum: The Impact of Autism Spectrum Disorders on Law and Policy, 45 VAL. U.L. REV. 415, 454 (2011).}\]
children's health—and autism is identified as a major research target. One year later, Congress approved the No Child Left Behind Act, which focuses on standardized testing and alternative testing for students with disabilities. The Combating Autism Act of 2006 (CAA) established the Interagency Autism Coordinating Committee (IACC), and the Combating Autism Act of 2011 ensures that programs established under the Act of 2006 continued for an additional three years. In 2013, President Obama, while celebrating World Autism Awareness Day, launched a 100 million dollar initiative, the Advancing Innovative Neuro-Technologies: BRAIN-Initiative, designed to revolutionize the understanding of the human brain and to help researchers find new ways to treat, cure, and prevent brain disorders. President Obama also signed the Autism Collaboration, Accountability, Research, Education and Support Act in August 2014, which authorized 1.3 billion dollars to fund ASD research, services, and support activities.

Since 1990, the United States has seen an impressive increase in the volume of litigation on NDD-related issues both at the state and federal level, likely buoyed by the lack of universal social services. Many areas of law have experienced a

---

123. Caruso, supra note 7, at 487.
128. This Act authorized 924 million dollars of federal investment in autism research, treatment and services. Id. § 399EE.
spike in NDD-related cases, including health care law, criminal law, family law, tort, and discrimination in the workplace. One of the most litigated legal issues involves the public education system and its regulation under Individual with Disability Education Act (IDEA).131 Two major U.S. Supreme Court ASD-related judgments deal with elements pertaining to IDEA.

During the 1970s children with severe disabilities—including ASD—were often excluded from attending school, and institutionalization was a common occurrence.132 Autism's etiology, with its supposedly low prevalence, was unclear, and it was not specifically mentioned as a disability category under the Education of the Handicapped Act of 1970.133 In 1975, the U.S. Congress specifically addressed the rights of students with disabilities in the Education for All Handicapped Children Act of 1975 (EAHCA), mandating that children with special needs receive an appropriate education in public school special programs and that they be placed in "the least restrictive environment."134 Congress amended the EAHCA in 1986 in order to create two new programs: 1) Early Intervention for Infants and Toddlers with Disabilities, serving children from birth to age three, and 2) the Pre-School Special Education Program for children from three to five.135 Once again, autism was not ex-


132. Dicker & Bennett, supra note 121, at 417.

133. Amendments with Respect to Handicapped and Neglected or Delinquent Children, Pub. L. No 91-230, § 105(a)(5), 84 Stat. 121, 123 (1970) (providing grants for the education of "handicapped children (including mentally retarded, hard of hearing, deaf, speech impaired, visually handicapped, seriously emotionally disturbed, crippled, or other health impaired children who by reason thereof require special education . . . .").


plicitly mentioned in the 1986 amendments. In 1990, Congress amended the act again, and autism was finally listed as one of the disorders under the definition of the term "children with disabilities." In subsequent reauthorization, the EAHCA was renamed the Individuals with Disabilities Education Act (IDEA). In the aftermath of the No Child Left Behind Act, Congress reauthorized and amended IDEA by the Individuals with Disabilities Education Improvement Act of 2004, which emphasizes accountability. The IDEA constitutes a cornerstone in the evolution of the right to education for NDD children by conditioning federal funding to states on compliance with IDEA statutory requirements. States must provide to every eligible child a “free and appropriate public education” (FAPE) by means of a uniquely tailored individualized education program (IEP) in a “least restrictive environment” that each family negotiates directly with schools. In this special education framework—that created a separate environment from the regular education system based on a deficit-oriented policy for “slow learners”—Winkelman v. Parma made some important improvements to the IEP complaint mechanism, and Endrew F. v. Douglas subsequently provided

136. See Dicker & Bennett, supra note 121, at 418 (describing the amendments).
140. See Wendy F. Hensel, Sharing the Short Bus: Eligibility and Identity Under the IDEA, 58 Hastings L.J. 1147, 1148, 1178 (2007) (arguing against rendering “special education” too broad lest school districts cannot manage programs).
new standards for evaluating if a state's education programs for disabled children are appropriate under the requirements of the IDEA.

The two judgments are intimately interconnected, as both critique the design of education for disabled pupils. Justice Kennedy, writing the opinion in Winkelman v. Parma, points out the issue underlying this case:

The question is whether parents, either on their own behalf or as representatives of the child, may proceed in court unrepresented by counsel though they are not trained or licensed as attorneys. Resolution of this issue requires us to examine and explain the provisions of IDEA to determine if it accords to parents rights of their own that can be vindicated in court proceedings, or alternatively, whether the Act allows them, in their status as parents, to represent their child in court proceedings.\textsuperscript{143}

By responding affirmatively to this question, the Supreme Court seems aware of the inequalities generated by this system and takes into account the burden placed on families struggling with the disabilities of their children and fighting to receive vital services. Families not only have to negotiate the IEP, but they also have to sustain court expenses and legal fees if their needs are not met.

Daniela Caruso clearly delineates the core problem of this way of organizing special education:

Autism was added to the IDEA in 1991 . . . The machine of due process, however, was set in motion, and parents began to use it to create from scratch what is now a rich culture of autism education . . . . After fifteen years of IDEA practice, with its principled emphasis on individualized educational plans and parental due process, special education litigation had already become what it is now: a myriad of unrelated, atomistic disputes, each focused by definition around only one child.\textsuperscript{144}

\textsuperscript{143} Id. at 520.

\textsuperscript{144} Caruso, \textit{supra} note 7, at 516–17; for further discussion of Winkelman and relevant policy implications, see Zirkel, \textit{Autism Litigation under the IDEA}, \textit{supra} note 131; Zirkel, \textit{The Autism Case Law}, \textit{supra} note 131;; Zisser & van Stone, \textit{supra} note 131.
The case *Endrew F. v. Douglas* addressed an even more delicate aspect of the special education system in the United States, namely the standards of programs schools must provide to comply with federal requirements. Before this early 2017 judgment, the Supreme Court established the standard in *Board of Education v. Rowley*, in which the Court held that federal law does not require maximization of a child’s learning potential, or that students with disabilities be brought up to the level of their peers.\(^{145}\) School Districts interpreted the IDEA standard differently, some of them much more generously,\(^{146}\) effectively amplifying the differences in services delivered across the United States. On a federal level, *Rowley* establishes that a school is not obligated to provide a student with ASD with the *best* program, but rather a program that is reasonably calculated to help the child achieve some educational benefit.\(^{147}\)

In *F. v. Douglas Cnty. Sch. Dist. Re-1*, doctors diagnosed plaintiff Endrew F. with ASD. When his progress in school stalled, his parents challenged his public school IEP and asked for private school tuition reimbursement. On appeal from the lower courts, the Tenth Circuit interpreted the *Rowley* decision, affirming “a child’s IEP is adequate as long as it is calculated to confer an educational benefit [that is] merely . . . more than *de minimis* . . . [and] Endrew’s IEP had been reason-

---


146. The Ninth Circuit concluded, based on post-*Rowley* amendments to the IDEA, that an IEP is required “to confer a ‘meaningful educational benefit.’” N.B. v. Hellgate Elementary Sch. Dist., 541 F.3d 1202, 1213 (9th Cir. 2008); see also Lester Aron, *Too Much or Not Enough: How Have the Circuit Courts Defined a Free Appropriate Public Education After Rowley?*, 39 SUFFOLK U.L. REV. 1, 7 (2005) (identifying circuit court splits); Dicker & Bennett, *supra* note 121, at 424 (comparing various circuits’ approaches). All of that does not prevent states from establishing increased standards. Thus, in *Burilovich ex rel. Burilovich v. Bd. of Education*, 208 F.3d 560, 565 (6th Cir 2000), Michigan’s statute added to the federal FAPE mandates by requiring that an IEP be designed to develop the “maximum potential” of a child. Dicker & Bennett, *supra* note 121, at 429.

ably calculated to enable [him] to make some progress.”

Although the Supreme Court, after granting certiori, was clearly dissatisfied with the Tenth Circuit’s narrow application of the Rowley standard, it did not agree with the plaintiffs’ vision that the IDEA “requires States to provide children with disabilities educational opportunities that are substantially equal to the opportunities afforded children without disabilities.” Justice Roberts, writing for the majority, remanded the case, calling for the application of a more generous standard, and affirming that schools have to offer IEPs reasonably calculated to enable children to make progress appropriate in light of a child’s circumstances. According to the Court, the language of the IDEA implies that educational programs should allow children with disabilities to progress from grade to grade if possible, but schools must always consider appropriate ambitions in light of a child’s circumstances: “[t]he goals may differ, but every child should have the chance to meet challenging objectives.” While the Court rejected the idea that disabled children should achieve academic success and attain self-sufficiency, it maintained that the judiciary should not substitute its own judgment for that of school officials. Instead, the Court considered a range of interpretations of the educational standards required for disability, ranging from “merely more than de minimis”, to IEPs able to give “appropriate ambitions in the light of a child’s circumstances,” through to maximizing the disabled child’s potential as it is equally done for non-disabled students.

While Endrew marked a major advancement and a clear victory for families and activists, the Court did not fully embrace the claims of the neurodiversity movement—that individuals with disabilities, and in this case people with different neurodevelopmental profiles, have the same rights, as a matter


149. Id. at 992 (emphasis added).

150. Id. at 999.

151. Id. at 1000.


of equality and dignity, to maximize their potential through the public educational system, no matter their pre-existing conditions. In a neurodiversity-oriented perspective, this parity of entitlement would be the only way to avoid differential citizenship and to create the pre-conditions necessary for social inclusion.

Section 15(1)154 of the Canadian Charter is one of the few constitutional provisions in the world— together with Section 9(3)155 of the 1996 South African Constitution—that recognizes mental health or disability as formal ground of discrimination. Despite the fact that the Canadian system strongly protects substantial equality, indirect impact, and redistribution, the Court in Ottawa unexpectedly decided a case on autism that disappointed civil and disability rights activists.156 In Auton v. British Columbia, before the Canadian Supreme Court, parents of autistic children challenged British Columbia’s failure to provide for a form of behavioral therapy called Applied Behavioral Analysis (ABA) for pre-school autistic children in the provincial public health plan.157 ABA was the first treatment identified by researchers as an effective therapeutic strategy for children with autism and is now the mostly commonly used therapeutic protocol for several diverse neurodevelopmental conditions.158 The Canadian Supreme Court refused to recog-

---


155. The Bill of Rights of South African Constitution of 1996, drafted after the end of the Apartheid era, recognizes in section 9(3) a large sets of grounds of discrimination. According to this provision: “The state may not unfairly discriminate directly or indirectly against anyone on one or more grounds, including race, gender, sex, pregnancy, marital status, ethnic or social origin, colour, sexual orientation, age, disability, religion, conscience, belief, culture, language and birth” (emphasis added).


158. This makes this Canadian case highly comparable with the US situation, as FAPE and IEP requirements have been repeatedly challenged in courts in order to compel schools to offer ABA as an essential educational
nize the public health care system’s obligation to provide this therapy for the following reasons: 1) the fact that ABA is often still considered a non-core service, despite growing medical and scientific support for the therapy’s efficacy. In this perspective, according to the Court: “the Canada Health Act and the relevant British Columbia legislation do not promise that any Canadian will receive funding for all medically required treatment”. Therefore, the exclusion of ABA for autistic children from non-core benefits does not amount to discrimination; and 2) the specific appropriate comparator group targeted by the Court in testing whether the equality was infringed.

In terms of the legal implications for neurodiversity efforts, especially where equality is concerned, the most significant portion of the Canadian decision is the second point. The type of discrimination potentially involved in this case is not a violation of the equal protection clause, but rather is experienced as an indirect impact concerning substantive equality. The Canadian Supreme Court has vast jurisprudence providing specific hermeneutical tests to solve cases involving indirect discrimination. As in any other constitutional system fighting systemic discrimination, one of the most complicated heuristic operations is determining the right comparator group to the one supposedly discriminated against by the disputed legislation. Equality is normally understood as a comparative concept requiring a claimant to point to some person program. U.S. standard of judicial interpretation of FAPE imposes restrictions and limits on what type of services children with ASD may receive. There has been much controversy surrounding whether parents are entitled to ABA. Courts, as in J.P. ex rel. Popson v. West Clark Cmty. Sch., 230 F. Supp. 2d 910 (S.D. Ind. 2002), often decide that a child is not necessarily entitled to an ABA program—even when parents prove the effectiveness of this program for their child—if the school provides an alternative program that gives the child some meaningful benefit. See Dicker & Bennett, supra note 121, at 426 (identifying J.P. ex rel. Popson as part of a trend).

who has been treated better as the foundation for a claim. According to the plaintiffs in the Auton case, the Court failed in this portion of the reasoning by not selecting the “non-disabled children and their parents, as well as adult persons with mental illness” as a comparator group. Justice McLachlin, on behalf of a unanimous court, rejected these groups with the following rationale: “the comparator group should mirror the characteristics of the claimant or claimant group relevant to the benefit or advantage sought, except for the personal characteristic related to the enumerated or analogous ground raised as the basis for discrimination.” Applying these criteria, the Court held that the appropriate comparator group is a non-disabled person or a person suffering a disability other than a mental disability (here autism) seeking or receiving funding for a non-core therapy important for his or her present and future health, which is emergent and only recently becoming recognized as medically required.

The judges in Ottawa concluded that there was no evidence of a comparator group that had received access to a non-core therapy. Consequently, they rejected the claim.


164. Id. ¶ 53.

165. Id. ¶ 55.

166. Although U.S. schools are not compelled to provide ABA if they have alternative programs, districts cannot refuse in principle to provide it. In Deal v. Hamilton Cnty. Bd. of Educ., 392 F.3d 840, 858 (6th Cir. 2004), the school district did not include ABA among its educational programs, as it had previously invested in an alternative approach. In rejecting the district’s determination, the court stated that IEPs should be based on the specific needs of the child, and that a school district cannot refuse to consider a specific method of treatment. See Dicker & Bennett, supra note 121, at 428 (“[E]ven though schools are not required by the IDEA to maximize a disabled child’s educational benefit, they must still consider any and all programs that may help a particular child learn . . . . [S]chool districts cannot have all-or-nothing policies; they cannot ban or only use ABA therapy.”). The judiciary displays a lack of understanding of ASD as a social and behavioral disability. For example, in Thompson R2/F Sch. District v. Luke P., 540 F.3d 1143, at 1151 (10th Cir. 2008), the Court required only that the student be
In the UK, the Supreme Court dealt with a similar NDD-related case in *A v. Essex County Council*. Parents of a twelve year old autistic boy who also had serious learning difficulties, a severe communication disorder, epilepsy and behavioral challenges, claimed that their son's educational and welfare rights under the 1998 H.R. Act and his right to education under Article 2 of the Protocol of the European Convention on Human Rights were denied as a result of the Essex County educational system's failure in establishing appropriate special school accommodation. Both parents and school staff deemed the boy unable to attend ordinary school classes due to the severity of his health and behavioral issues. However, Essex County school authorities were unable to find a proper permanent placement over a period of eighteen months, in which the boy's behavioral and cognitive situation deteriorated. Unfortunately, the facts were relatively clear. Although the worsening of symptoms were due to: "(ii) lack of sensory stimulation, boredom, and lack of meaningful occupation; (ii) inability to clearly communicate his needs and be clearly understood by those around him,"—things reasonably connected to the blunt interruption of school activities—the Court established that "Essex were doing their utmost to have [the complainant] properly appraised and thereafter did their utmost to arrange residential care, for which they paid." The court held that the plaintiff was not denied the "very essence" of his right to education. This case shows the potential severity of impairments and highlights the vulnerability of ASD individuals and their families.
In 2014 and 2016, the Spanish and German Constitutional Tribunals also decided two cases on issues related to NDDs and the public education system.\textsuperscript{174} These cases demonstrate how challenging the governance of special education services can be when related to children with ASD or Asperger’s syndrome. Although the two cases are facially different—as one challenges the decision to assign an ASD child to the special education system, while the other claims a delay in special education services—they demonstrate the deep concerns of families and caregivers, as well as the hurdles and difficulties they encounter in navigating the public school system.

In 2016, the German Constitutional Tribunal decided a neurodiversity-related case involving public school special services.\textsuperscript{175} A family of a child diagnosed with Asperger’s syndrome claimed violation of several fundamental constitutional rights, namely Articles 2 and 3 of the German Constitution, for delay of special educational services and extra teaching hours conducted by a specialized teacher at the beginning of the school year.\textsuperscript{176} The German Constitutional Court struck down the appeal, holding that the plaintiffs failed to establish meaningful concerns about the delay of special education services, and that the Administrative Tribunal fairly applied both the substantial and procedural law on social service, public schooling and mental disability.\textsuperscript{177} Perhaps most significantly, this case demonstrates that similar litigation is emerging in the United States and within Europe.

The Spanish Constitutional Tribunal, with its 2014 \textit{Recurso de Amparo}, was the second European constitutional jurisdiction that issued a decision in autism-related litigation.\textsuperscript{178} In this case, the parents of an autistic child claimed a violation of his

\textit{a Neurodiversity Spectrum Statement}, 17 TIZARD LEARNING DISABILITY REV. 36 (2012) (analyzing laws, policies and reforms focused on special educational needs and equality in England and to suggest a Neurodiversity spectrum statement).


\textsuperscript{175} BVerfG, 1 BvR 3514/14, Feb. 14, 2016, http://www.bverfg.de/e/rk20160214_1bvr351414.html (Ger.).

\textsuperscript{176} Id. ¶¶ 2, 5.

\textsuperscript{177} Id. ¶¶ 6–8.

\textsuperscript{178} S.T.C., Jan. 27, 2014 (S.T.C., No. 48, p. 47) (Spain).
right to public education. Their child was placed in a special education facility separated from ordinary public education classes, and his parents claimed the decision constituted discrimination. They claimed a violation of the constitutional principle of equality due to the unequal treatment inflicted upon their son as well as a violation of moral integrity and dignity. Although the Constitutional Tribunal recognized that Spain had ratified the UN Convention on the Right of Persons with Disabilities, and that Article 74.1 of the Ley Orgánica 2/2006 affirms that education for children with special needs must follow the principles of normalization, inclusion, non-discrimination, and equality in access to the educational system, they nevertheless held that enrollment in special educational facilities is legally and constitutionally permitted if its aim is protection of the education of children with special needs. The Court also held that the school authorities' justification for this decision must be reasonable, and inclusion in an ordinary public education school must be pursued if the special child needs are not disproportionate. On the facts of the instant case, the Constitutional Tribunal ruled that the school administration did not act on disproportionate or unreasonable grounds in deciding to place the child in a special educational facility. From a comparative perspective, this case matches similar interpretations in the United States and UK on special educational arrangements. In these jurisdictions special educational programs where children with special needs are placed in separated school environments are perfectly legal. The European Committee on Social Rights of the Council of Europe (ECSR) has decided two cases related to NDDs and public education issues. The European Social Charter

179. Id. at 48–51.
180. Id. at 51.
181. Id. at 57–58.
182. Id. at 61–62.
183. Id.
184. The remaining question—involving a more thorough pedagogical analysis that goes beyond our purposes here—is whether special education programs have positive effects in making NDD children acquiring better academic or behavioral results.
International Law and Politics (Vol. 51:69)

European Social Charter (ESC), an international treaty supplementing the European Convention on Human Rights in the field of economic and social rights, entrenches fundamental rights and freedoms and establishes a supervisory mechanism based on a system of collective complaints and national reports, guaranteeing their respect by state Parties.\textsuperscript{186} After amendment of the Charter in 1996, it is now ratified by forty three of the forty seven Council of Europe Member States.\textsuperscript{187} One of the major elements of the ESC was the establishment of a specific judicial body, the ECSR, whose mission is determining if state parties are in conformity in law and in practice with the provisions of the ESC.\textsuperscript{188} The ECSR also has the power to adopt conclusions and issue decisions regarding collective complaints for state violations of the Social Charter.\textsuperscript{189} Under a protocol opened for signature in 1995, which entered into force in 1998, complaints of violations of the Charter may be lodged with the European Committee of Social Rights.\textsuperscript{190}

One of the first comprehensive judicial evaluations of state public special education programs emerged in \textit{International Association Autism-Europe (IAAE) v. France},\textsuperscript{191} in which the ECSR decided a case brought by Autism Europe against France for violation of Articles 15 and 17 of the ESC, as well as violation of the principle of non-discrimination entrenched by Article E of the ESC. The ECSR concluded that France failed to meet its obligations to people with autism under the ESC on several bases. First, the ESC, which requires that states guarantee the right to education to people with disabilities, clearly includes both children and adults with autism. Second, notwithstanding debates about the number of persons concerned.

\begin{thebibliography}{9}
\item \textsuperscript{186} European Social Charter, Oct. 18, 1961, E.T.S. No. 35.
\item \textsuperscript{188} Additional Protocol, \textit{supra} note 187, art. 5 (establishing a Committee of Independent Experts).
\item \textsuperscript{189} \textit{Id.} art. 8.
\item \textsuperscript{190} \textit{Id.} arts. 5–8.
\end{thebibliography}
and the relevant strategies required, France failed to achieve sufficient progress in advancing the provision of education for persons with autism. Third, France still used a more restrictive definition of autism than the one adopted by the World Health Organization (WHO) in violation of the ESC. Finally, the proportion of children with autism receiving education in either general or special schools was much lower than other children, whether or not disabled. The decision highlights that autistic people are an excluded group within the excluded. The decision also notes a chronic shortage of care and support for autistic adults. In addition, since Article E prohibits not only direct discrimination but also indirect discrimination arising when people with autism are given “inappropriate treatment,” France must use available resources to meet the needs of people with autism and their families.

This decision has been used extensively as a precedent in subsequent cases before the ECSR.

France was again challenged before the ECSR in European Action of the Disabled (AEH) v. France in 2012 for violation of Article 15(1) in connection to Article E. In this case, the Committee found that France recognizes autism as a disability under the old laws of 1975 and 1996, which remain in force today. It also observed that under this legislation, the state prioritizes educating children and adolescents with autism in mainstream schools. In 2005, the State established the right for all children with disabilities, including those with autism, to be enrolled in a mainstream school, and provided alternatives such as individual schooling in mainstream classes with

192. Id. ¶¶ 47-54.
193. Id. ¶ 54.
194. See id. (explaining that indirect discrimination is demonstrated by the proportion of autistic children being educated in either mainstream or special schools, which was significantly lower than that of other children with or without a disability).
197. Id. ¶ 82.
the help of a school assistant. Provision was also made for teaching in separate classes at mainstream schools that a collective school assistant runs. These school integration classes occur at the early childhood and primary levels, and lower and upper secondary levels.\textsuperscript{198} In its decision, the ECSR affirmed:

\[\text{T}\]he lack of opportunities for special classes to integrate with mainstream classes; the scarcity of joint lessons with mainstream classes; the fact that children with autism do not have their recreational and lunch breaks at the same time as other children in the school; the isolation of classrooms for special classes from other buildings and overcrowding of these classes; the fact that children with all types of disability are mixed together in special classes; a high rotation of teaching teams, conducive neither to regular teaching nor to the consistent implementation of suitable teaching programmes; the lack of specific training for teachers about autism, with the result both that they are reluctant to allow children with autism to join their class and that unsuitable teaching methods may be adopted and; the fact that teachers who are concerned about the proper integration of children and adolescents with autism are left to arrange their own training on autism.\textsuperscript{199}

Although French legislation and policy did progress since the first judgment in 2003, the ECSR still identified a number of violations of the Social Charter such as: a) "the limited funds in the state’s social budget for the education of children and adolescents with autism indirectly disadvantages these persons with disabilities,"\textsuperscript{200} and b) "families have no other choice than to leave the national territory in order to educate their children with autism in a specialised school, which constitutes a direct discrimination against them . . ."\textsuperscript{201}

\textsuperscript{198} The ECSR would point out that the number of children with autism in school decreases significantly: 87% of children with attend primary school, 11% lower secondary school and 1.2% upper secondary school. \textit{Id.} ¶ 83.

\textsuperscript{199} \textit{Id.} ¶ 90.

\textsuperscript{200} \textit{Id.} ¶ 145.

\textsuperscript{201} \textit{Id.} Conclusion.
A final significant health care and neurodiversity-related litigation is *Bruesewitz v. Wyeth LLC*, in which autism drew the U.S. Supreme Court's attention for different reasons: the very controversial establishment of a cause-effect connection between vaccines and the onset of these conditions in American children, and the consequential damage compensation claimed by thousands of plaintiffs in the United States. In the face of converging scientific evidence from dozens of epidemiologic studies that autism is not an immune-mediated disease, the officially discredited theories postulating a connection between MMR vaccines and NDDs fueled the global anti-vaccine movement. In 1986, the U.S. legislature established the National Vaccine Injury Compensation Program by adopting the National Childhood Vaccine Injury Act (NCVIA). With this act, Congress assigned jurisdiction to the Court of Federal Claims for implementation of a no-fault compensation program in the hopes of stabilizing a vaccine market adversely affected by an increase in vaccine-related tort litigation. This consequently created the Omnibus Autism Proceeding (OAP), a mechanism to adjudicate thousands of peti-

---


tions filed by parents and families under the legal framework.\textsuperscript{206}

In \textit{Bruesewitz v. Wyeth LLC}, the Supreme Court held that the NCVIA “pre-empts all design-defect claims against vaccine manufacturers brought by plaintiffs who seek compensation for injury or death caused by vaccine side effects.” In so doing, the Court closed the door on thousands of claims by parents alleging a link between vaccines and childhood autism.\textsuperscript{207} With Justices Sotomayor and Ginsburg dissenting, the Court dealt with the highly complex structure of vaccine regulation in relation to the NCVIA. Among several issues decided by the Court, one of the most important was that the licenses issued for vaccines require warnings, directions, and manufacturing methods.\textsuperscript{208} Although the FDA does not regulate design defects, the NCVIA “micromanages manufacturers” except on “how to evaluate competing designs.”\textsuperscript{209} According to the majority, the Act focuses on the benefits of design-defect torts by encouraging improvement of vaccine design by federal agencies, and by compensating victims.\textsuperscript{210} The Court concluded that Congress’s omission of design-defect liability reflects a “sensible choice to leave complex epidemiological judgments” to experts instead of jurors.\textsuperscript{211} Although extensive OAP litigation confirmed that no correlation between vaccines and NDDs has been determined, “unavoidable” side effects might occur.\textsuperscript{212} Instead of facilitating extremely complex and highly expensive vaccine design-defect litigation to address these side effects, NCVIA establishes a victim compensation mechanism based on a manufacturer’s fund.

This third NDD-related case decided by the U.S. Supreme Court, and the thousands of applications filed under the OAP, evince families’ deep concerns and profound difficulties dealing with these challenges. Compensation might help them deal with a disability health and social care system that pro-

\begin{itemize}
\item \textsuperscript{206} Keelan & Wilson, \textit{supra} note 202, at 2016; see also Paul A. Offit, \textit{Vaccines and Autism Revisited—The Hannah Poling Case}, 358 \textit{NEW ENG. J. MED.} 2089, 2090 (2008) (explaining implications of the NCVIA framework).
\item \textsuperscript{207} Bruesewitz \textit{v. Wyeth LLC}, 562 U.S. 223, 243 (2011).
\item \textsuperscript{208} Id. at 237.
\item \textsuperscript{209} Id. at 238.
\item \textsuperscript{210} Id.
\item \textsuperscript{211} Id. at 239.
\item \textsuperscript{212} Id. at 230.
\end{itemize}
vides insufficient support. Interestingly, these cases have a common premise: the search for a unique cause of NDDs. This constitutes one of the major conceptual and agenda-related clashes between neurodiversity self-advocates and other stakeholders, including some parents’ organizations.

B. Physical Liberty, Competency, and Criminal Law

Criminal law, physical liberty, and competency are other extremely sensitive neurodiversity-related areas. In the last decade, European jurisdictions dealt with several such cases both at national and supranational level. The 2014 UK Supreme Court case *P v. Cheshire and Chester Council et al.; P and Q v. Surrey County Council*213 dealt with a sensitive aspect of neurodiversity-related litigation, namely “the criteria for judging whether the living arrangements made for a mentally incapacitated [or autistic] person amount to a deprivation of liberty. If they do,” the Court maintained, “then the deprivation has to be authorised, either by a court or by the procedures known as the deprivation of liberty safeguards . . . .”214 The Court continued: “human rights are for everyone, including the most disabled members of our community, and . . . those rights include the same right to liberty as has everyone else.”215 The complainants were two sisters, MEG and MIG, with a history of neurodevelopmental issues, such as learning disabilities, problems with hearing, difficulties in communication, autistic traits, and challenging behaviors. The other complainant, P., had cerebral palsy and Down syndrome requiring twenty-four hour care.216 The two sisters were placed in foster families after experiencing abuse and ill-treatment in their original family.217 They received differently organized living arrangements that the Court of Protection ultimately determined did not amount to a deprivation of liberty.218 P., after his mother was no longer able to be the principal caregiver, was placed in Z. house, in which he received living arrange-

214. *Id.* at [1].
215. *Id.*
216. *Id.* at [16].
217. *Id.* at [12].
218. *Id.* at [14]–[15].

ments that the Court of Appeal decided did not constitute a deprivation of liberty.

The three plaintiffs were individuals who needed intense support and control. As the UK Supreme Court meaningfully states:

\[
\text{[P]eople with disabilities, both mental and physical, have the same human rights as the rest of the human race. It may be that those rights have sometimes to be limited or restricted because of their disabilities, but the starting point should be the same as that for everyone else. This flows inexorably from the universal character of human rights, founded on the inherent dignity of all human beings, and is confirmed in the United Nations Convention on the Rights of Persons with Disabilities. Far from disability entitling the state to deny such people human rights: rather it places upon the state (and upon others) the duty to make reasonable accommodation...}^{219}
\]

This case is "a good illustration of the sort of benevolent living arrangements which many might find difficult to characterise as a deprivation of liberty."^{220} In such situations, the Court must decide if the complainant's living arrangements constitute a deprivation of *physical* liberty. In discussing the right to physical liberty in the context of NDDs, the Court wrote,

This is not a right to do or to go where one pleases. It is a more focused right, not to be deprived of that physical liberty. But... what it means to be deprived of liberty must be the same for everyone, whether or not they have physical or mental disabilities. If it would be a deprivation of my liberty to be obliged to live in a particular place, subject to constant monitoring and control, only allowed out with close supervision, and unable to move away without permission even if such an opportunity became available, then it must also be a deprivation of the liberty of a disabled person. The fact that my living arrangements are comfortable, and indeed make my life as enjoyable as

219. *Id.* at [45].
220. *Id.* at [10].
it could possibly be, should make no difference. A gilded cage is still a cage."  

The Court’s arguments touch a sensitive area of the idea of equality, rejecting the idea of “relative normality” in which disabled people’s life situation must be compared to another disabled individual’s life. Instead, deprivation of physical liberty occurs when a person is under continuous supervision and control and is not free to leave if he expresses the desire to do so, regardless of ability. In conclusion, the Court states that: “[i]f the acid test is whether a person is under the complete supervision and control of those caring for her and is not free to leave the place where she lives . . . both MIG and MEG are being deprived of their liberty.” In addition: “[b]ecause of the extreme vulnerability of people like P, MIG and MEG . . . [t]hey need a periodic independent check on whether the arrangements made for them are in their best interests . . . [t]hey are a recognition of their equal dignity and status as human beings like the rest of us.”

This legal and ethical problem of what constitutes deprivation of liberty in the context of NDDs is widely discussed in other legal systems in which “least restrictive means” and periodic reporting and review of guardianship requirements are becoming the more common approach. This transformative and liberating approach comes after decades during which courts deemed people with mental impairments wholly incapable of retaining any legal rights or autonomy. The right of

221. Id. at [46].
222. Id. at [47]. The National Autistic Society and Mind, intervening in this case, expressed a favorable opinion. See id. at [50] (noting the National Autistic Society’s adoption of a similar test, rejecting the notion of “relative normality”).
223. Id. at [54]. The Court continues: “that deprivation is the responsibility of the state. Similar constraints would not necessarily amount to a deprivation of liberty for the purpose of article 5 if imposed by parents in the exercise of their ordinary parental responsibilities and outside the legal framework governing state intervention in the lives of children or people who lack the capacity to make their own decisions.” Id.
224. Id. at [57].
226. Cohen et al., supra note 51, at 408; see also Robert D. Dinerstein, Implementing Legal Capacity under Article 12 of the UN Convention on the Rights of Persons with Disabilities: The Difficult Road from Guardianship to Supported Deci-
physical liberty is one of the pivotal fundamental rights, and as such, it is protected in the UK by human rights statutes, by Article 5 of the European Convention on Human Rights, and by the 2007 Mental Capacity Act (MCA). Although "[t]hroughout the 19th century it was assumed that persons of unsound mind (then known as either 'lunatics' or 'idiots') should be kept in some form of confinement and reformers concentrated upon providing more and better institutions where they could live," deprivation of liberty is now impermissible under the MCA, save in three circumstances: 1) when authorized by the Court of Protection; 2) when authorized under the procedures provided for in Schedule A1, which relates only to deprivations in hospitals and in care homes falling within the meaning of the Care Standards Act 2000; and 3) if deprivation is necessary in order to give life sustaining treatment or to prevent a serious deterioration of the person's condition while a case is pending before a court.

The European Court on Human Rights (ECtHR) decided similarly significant cases. That Court rules on individual or state applications alleging violations of the European Convention on Human Rights. Since 1998 it has sat as a full-time court to which individuals can apply directly. In this context, the ECtHR decided an extensive number of cases on mental health issues and civil liberties.

In H.L. v. United Kingdom, the ECtHR dealt with the allegations from an autistic applicant who claimed to have been...
detained in a psychiatric institution as an “informal patient.” The applicant alleged that the review procedures available regarding the legality of his detention, as well as his treatment in that institution, did not satisfy the requirements of the Convention.231 Before reaching the ECtHR, the UK Court of Appeal in 1997 and by the House of Lords in 1998 heard the case.232 During these procedures, it was established that the applicant’s detention in a psychiatric facility was justified under a common law doctrine of necessity according to the Mental Health Act of 1983.233 In its decision, the ECtHR established that the “absence of procedural safeguards fails to protect against arbitrary deprivations of liberty on grounds of necessity and, consequently, to comply with the essential purpose of Art. 5 [of the Convention].”234 With this, the Court held that the UK authorities, in deciding the applicant’s stay in the facility without procedural safeguards, violated the Convention.

One year later, the ECtHR heard another similar case, Storck v. Germany, in which an applicant showing signs of autism challenged her repeated placement in a psychiatric institution.235 The applicant (who is currently 100% disabled236) has spent almost twenty years of her life in psychiatric institutions and other hospitals.237 The Court also uncovered evidence of serious conflicts between the applicant and her parents, and that her father believed her to be suffering from psychosis. Judges of the ECtHR, establishing the background circumstances, affirmed that:

The applicant—who by that time had attained the age of majority—had not been placed under guardianship, had never signed a declaration that she had consented to her placement in the institution, and there had been no judicial decision authorising her detention in a psychiatric hospital. The private clinic

232. Id. ¶¶ 33-37.
233. Id. ¶ 12.
234. Id. ¶ 124.
236. Id. ¶ 12.
237. Id.
was not entitled to detain patients who were to be kept in accordance with the Act of the Land of Bremen on the detention of mentally insane persons, mentally deficient persons and drug addicts . . . .

[T]he police brought the applicant back to the clinic by force after she had attempted to escape.238

Due to the long stay at the clinic, the applicant was unable to maintain social connections with persons outside the clinic, and she developed additional physical pathologies that required different medical interventions. Significantly, subsequent psychiatric evaluations assessed that the applicant: "at no point in time suffered from a schizophrenia-type psychosis . . . and that her intemperate behaviour had resulted from conflicts with her family."239 An additional evaluation determined that: "the applicant had never suffered from an early onset of schizophrenia . . . ."240 According to the Court, this case entailed "deprivation of liberty without a judicial decision, [and the] absence of a legal basis for the detention . . . ."241

After national judicial remedies proved fruitless and the Federal Constitutional Tribunal refused to accept the applicant's constitutional complaint, the applicant brought her case before the ECtHR. The Court considered the following elements pertinent: the applicant,

[T]ried on several occasions to escape. She had to be shackled in order to prevent her from absconding and brought back to the clinic by the police when she managed to escape on one occasion. Under these circumstances, the Court is unable to discern any factual basis for the assumption that the applicant—presuming that she had the capacity to consent—agreed to her continued stay in the clinic. In the alternative, assuming that the applicant was no longer capable of consenting following her treatment with strong medication, she cannot in any event be considered to have validly agreed.242

238. Id. ¶ 15.
239. Id. ¶ 22.
240. Id. ¶ 23.
241. Id.
242. Id. ¶ 76.
This case involves the problem of if and how vulnerable individuals have the right to express their wish to stay or leave psychiatric facilities, and whether this expression of will and its refusal determines that their right to physical liberty has been infringed. Here, the Court found a violation of Article 5(1) of the European Convention for some of the plaintiff’s periods of confinement.

The ECtHR case Nelissen v. The Netherlands centered on compulsive and repetitive behaviors resulting in restricted interests, a common symptom of NDDs. The phenomenology of these behaviors can worsen when individuals with neurodiverse conditions suffer unfamiliar or stressful events that trigger disruptive and deregulated reactions. As noted above, serious problems may arise when the law considers the object of the restrictive interests, or the circumstances related to them, a criminal offence. In the course of criminal proceedings in the Netherlands, it emerged that the applicant in Nelissen v. The Netherlands was a pathological collector of devotional obituary cards for the recently deceased. As the ECtHR later reported, “whenever possible [he] would gatecrash funerals in the Maastricht area.” In 2003, the applicant went to the home of the victim and demanded that she give him an obituary card of her sister who had recently died. On being met with a refusal, he forced his way into the victim’s home, causing her to fall. The applicant “had met [the victim] at her sister’s funeral, from which [the applicant] had been turned away by the undertaker. At the time of his visit to [the victim’s] home, he had been under the mistaken impression that he had propitiated the victim by sending her a blessed candle beforehand.” In pre-trial detention, the applicant underwent a psychiatric assessment: “[e]xaminations of the applicant’s mental state by two psychiatrists and a psychologist showed that the applicant was . . . suffering from a mental im-

243. Id. ¶ 112.
244. See Newman & Ghaziuddin, supra note 79, at 1851 (“Persons with Asperger syndrome have been described as lacking in conscience, which increases their risk for committing crime. However . . . aggressive behavior is not one of its defining or discriminating features.”).
246. Id. ¶ 6.
247. Id. ¶ 7.
pairment. The report of the psychiatrists diagnosed it as 'a paranoid form of schizophrenia in which autistiform characteristics are unusually prominent and . . . an obsessive compulsive personality disorder.' 248 Moreover:

[B]oth the psychiatrists’ report and that of the psychologist considered it very likely that the applicant would reoffend if given the chance. Both reports mentioned the applicant’s lack of awareness of his condition, which made him difficult to treat. They concurred in recommending that the applicant undergo compulsory treatment in a closed setting. In addition, a psychologist attached to a mental hospital where the applicant had been held previously, under the Psychiatric Hospitals (Compulsory Admission) Act . . . gave evidence in open court to the effect that the applicant had absconded before and would resist ‘all possible forms of interference in his personal life’ if he could.249

In 2004, the ‘s-Hertogenbosch Court of Appeal in the Netherlands sentenced the applicant to seven months imprisonment and confinement in a custodial clinic.250 Before the ECtHR the applicant challenged the excessive length of deprivation of liberty calculation, the pre-trial detention, prison detention, and confinement in a custodial clinic. While the ECtHR found a violation of Article 5.1 of the ECHR,251 the statement by the Government of the Netherlands is significant:

[T]he applicant suffered from Asperger’s syndrome and obsessive compulsive disorder. Although Asperger’s syndrome was a congenital condition and incurable, the reports on the applicant’s mental state suggested that his behavioural problems could be made manageable. This was important, since, as was commonly the case with Asperger’s syndrome, any

248. Id. ¶ 9.
249. Id.
250. Id. ¶ 3.
251. See id. ¶ 60 (“[T]he TBS order [confinement in a custodial clinic] imposed on the applicant became final and enforceable on 14 February 2006. The applicant was admitted to a custodial clinic on 20 March 2007, that is one year, one month and six days later. Such a time lapse cannot be considered acceptable.”).
loss of structure in the applicant’s life could lead to aggression. This in turn could lead to a heightened risk of the applicant’s reoffending. Given the nature of the applicant’s disorders, the applicant’s detention had been reviewed with sufficient frequency. Specialists may easily recognize here common NDD traits. Given the commonality of some of these traits, three elements should be considered: 1) the importance of expanding wide social services for individuals with NDDs so that they can manage potentially disruptive behaviors; 2) creating greater awareness and understanding of NDD possible behavioral phenomenology within the legal system; 253 and 3) introducing some sort of flexible legal approach able to fairly evaluate when specific offences are clearly connected to NDD clinical conditions. These three elements are reasonably in line with an approach encompassing neurodiversity and brain equality.

C. Discrimination at the Workplace

Working environments can create consistent difficulties for individuals with NDDs and discrimination at the workplace generates sensitive neurodiversity-related cases. 254 On this subject, the Canadian Human Rights Tribunal decided Dawson v. Canada Post Corporation. Section 15 of the Charter of Rights and Freedoms (the Charter) establishes a wide equality and non-discrimination framework binding the federal and provin-

252. Id. ¶ 65.

253. The need to expand scientific knowledge within the legal system, especially when it comes to brain functioning and related behaviors is explained by David L. Faigman et al., A Matter of Fit: The Law of Discrimination and the Science of Implicit Bias, 59 Hastings L.J. 1389, at 1432 (2008); see also Jones et al., supra note 52, at 3-4 (discussing the use of neurological evidence in court); Sapolsky, supra note 53, at 580; Owen D. Jones, Law, Evolution and the Brain: Applications and Open Questions, 359 Phil. Transactions Royal Soc. B 1697 (2004) (advocating the use of neuroscience and evolutionary biology to make the legal system more effective); Sapolsky, supra note 53, at 1793 (surveying literature on neuroscience and the criminal justice system).

cial governments. The Canadian Human Rights Act is the most important federal statute implementing the principles of the Charter.256 Moreover, all the provinces have enacted Human Rights Codes to protect from unfair discrimination by private individuals and associations and provide remedies.257 In most provinces, Human Rights Commissions have been established to administer human rights legislation.258 Commissions receive complaints, investigate and seek to settle them. A commission, after investigation, may refer to the relevant Human Rights Tribunal for adjudication of claim. The jurisprudence developed by the Canadian Supreme Court in respect of Section 15 of the Charter is important to the determination of discrimination complaints lodged under the Canadian Human Rights Act as well as the various Human Rights Codes.259

_Dawson v. Canada Post_ is of utmost importance for: 1) the high level of analysis provided by the expert testimony and by the complainant herself; and 2) the type of remedies ordered by the Tribunal that accord with the requirements of a neurodiverse approach. The complainant, a person with ASD, alleged that the respondent discriminated against her on the basis of her disability. She alleged suffering adverse differential treatment by reason of her disability due to respondent treating her in an intolerant and paternalistic way and negligently failing to provide a harassment-free work environment—all in breach of Section 7 of the Canadian Human Rights Act.260 On top of these allegations, the complainant claimed that the structure of the hearing procedure before the Tribunal did not meet the needs and characteristics of an ASD individual.

255. See O'Regan & Friedman, _supra_ note 161, at 476 (discussing section 15 of the Charter of Rights and Freedoms).


258. _Id._


An expert witness emphasized that the "level of impairment that appears may be completely discrepant with the actual level of intelligence for autistic people . . . the idea of curing autism is meaningless." Moreover, "what characterizes autism is that the specific part of the brain which is in charge of social activities is broken in an autistic brain . . . [A]utistic people are vulnerable to the malice of their peers and that they are stigmatized in society. Autistic people are not violent by definition." In contrast, the expert testified that "people are intimately convinced that most autistic individuals are violent, even if it is not true. In general, non-autistic people have a poor understanding of autistic people." More importantly, among several core statements about the implications of neurocognitive differences on behaviors, the expert stated: "autistic people live in fact in a world that is not made for them." The conclusions of the Tribunal were thus striking:

An autistic person should expect that his workplace be free of any misperception or misconception about his condition. It goes to the right of autistic individuals to be treated equally, with dignity and respect, free of any discrimination or harassment related to their condition. In this respect, in a society where human rights are paramount, an employer has the duty to dispel such misconception or misperception about such individuals . . . To discriminate on the basis of somebody’s physical appearance or social behavior might be one of the cruelest forms of discrimination.

As a result, the Tribunal ordered that the respondent: a) work for a period of one year with the Commission to modify its existing policies to conform with human rights law of Canada; and b) retain, in consultation with the Commission, appropriate persons to conduct workplace equity, accommodation and sensitivity training for managers and staff, notably in relation to autism.

261. Id. ¶ 102.
262. Id. ¶ 103–04.
263. Id. ¶ 104.
264. Id. ¶ 108.
265. Id. ¶¶ 243, 246.
VI. Conclusion

This study provides some preliminary results. First, judicial analysis shows that there are substantial neurodiversity-related cases brought before European and North American courts. The litigation confirms the importance of the legal implications of the concept of neurodiversity. Second, the case law analyzed presents a preponderance of social equality claims. This matches some of the priorities of NDD-rights agendas. Comparatively, these social justice claims are highly similar across legal cultures. In this perspective, according to the case law analyzed, litigants demand wider public education services in the United States, the UK, Spain, Germany and France. This represents the most urgent equality request of neurodiverse people and their families. Individuals also seek an expansion of health care and social services in Canada, the United States, and the UK. European courts, in connection with the ECtHR, address neurodiversity-related litigation involving criminal law, physical liberty, and competency more than U.S. and Canadian courts. Courts that decide NDD-related cases involving the limitation of physical liberty are aware that thorough safeguards are required when health or social care bodies determine living arrangements for neurodiverse individuals.

From a legal perspective, the concept of neurodiversity fuels the debate around complex, open questions of modern constitutionalism. In particular, the hypothesis that brain-based diversities should receive wider constitutional recognition centralizes discussion on the doctrine of constitutional equality. First, it stresses the problem of ensuring pluralism while preventing the fragmentation of legal systems into countless differential norms. This implies considering if and to what extent a certain degree of disparate impact, without clear discriminatory intent, is an unavoidable effect of the social cohesion pursued by the universality and generality of the law.266 Second, the legal recognition of neurodiversity falls into the dichotomy of essentializing or de-essentializing identity—anti-subordination versus anti-classification—which triggers the prob-

---

266. See Kenji Yoshino, The New Equal Protection, 124 HARV. L. REV. 747, 774 (2011) (“[In the Smith [494 U.S. 872, 885 (1990)] decision . . . the Court underscores that exempting religious practices from laws of general applicability would permit every citizen 'to become a law unto himself.'”).
lem of harmonizing different ideologies about the public relevance of elements belonging to the private sphere. This debate revolves around the problem of how to harmonize two diverging needs: containing the proliferation of protected groups, and, at the same time, recognizing group-based diversity claims. Contemporary centrifugal forces undermining the unity of social and political bodies in Western countries influence this polarization. Although analysts raise concerns about the group-rights framework, and warn against subgrouping emphasis (balkanization), the worsening of the conditions of unprivileged, excluded, and neglected groups is nonetheless a socio-economic reality. In conclusion, the concept of neurodiversity fuels the discussion on how to determine legally precise criteria of group classification. In the field of different neurological profiles this is even more problematic. A formal group classification based on neurobiological traits would have a highly controversial bio-political meaning.

It is important to note, however, that it is unlikely that the definition of neurodiverse individuals as a minority group would directly provide them with any extra-legal protection, as the status of collective rights is controversial or poorly imple-


mented in most national legal systems. At the same time, international protection of minority rights still encounters resistance and difficulty since international conventions are not universally ratified, not directly enforced, and international jurisdictions have little power to force states to abide by them.

The legal implications of neurodiversity must therefore be contextualized in contemporary theories and practices of the constitutional equality. Over the past decades, the traditional group-based equality jurisprudence in the United States was limited due to growing concerns about dynamics of social division, also known as pluralism anxiety. Therefore, social fragmentation is governed by limiting the range of legally relevant diversity-group equality claims. Instead, U.S. scholars report a move toward liberty-based dignity claims under the due process guarantees of the Fifth and Fourteenth Amendments. In this context, despite consistent advancement both in legisla-


271. See Palermo & Woelk, supra note 95, at 10 (2003) (surveying the history and status of minority rights in international law).

272. Yoshino, supra note 266, at 774 (discussing of two types of pluralism anxiety: proliferation of classifications and proliferation of classes within classifications. Both “operate to curtail the equal protection jurisprudence.”) In *City of Cleburne v. Cleburne Living*, 473 U.S. 432 (1985), “[t]he Court declined to grant classifications discriminating against individuals with mental retardation heightened scrutiny.” Id. at 758. Moreover, in * Bd. of Trustees of Univ. of Alabama v. Garrett*, 531 U.S. 356 (2001), “Chief Justice Rehnquist [affirmed] that classifications based on disability drew only rational basis review under Section 1 of the Fourteenth Amendment.” Id. at 773.

tion and jurisprudence in the field of ASD and NDDs, in the United States the right to education does not yet maximize a disabled child's potential as it does for non-disabled students.

The Canadian constitutional ethos traditionally includes pluralism and group-based equality.\textsuperscript{274} The inclusion of people with mental disability as a designated group for formal constitutional protection under Section 15 of the Canadian Charter was viewed as a triumph of disability advocacy. However, as commentators point out, despite its symbolic power and several cases decided by the Supreme Court, the substantial impact of this framework is not very encouraging.\textsuperscript{275} Over the last two decades, Canadian specialists reported persistent difficulty in blending the medical model of mental disability with a more inclusive social model of disability.\textsuperscript{276}

In Europe, the multilayer configuration of the EU, the ECHR, and domestic constitutional systems, produces a highly complex legal structure. Criteria for determining disability as well as anti-discrimination mechanisms provided by national legislation differ widely throughout EU Member states.\textsuperscript{277} As a consequence, it is difficult to paint a comprehensive picture.

\textsuperscript{1694} (2008) (highlighting the interpretation of individual rights guarantees \textit{vis-à-vis} abortion).

\textsuperscript{274}. \textit{See} Hamilton & Shea, \textit{supra} note 157, at 127 (analyzing five Supreme Court of Canada judges' approaches to equality).


However, despite a proliferation of national strategy plans, special legislation and policy documents pleading for an inclusive concept of disability, especially in the realm of mental disability and NDDs, experts suggest that disability substantive discrimination and exclusion on the basis of disability still remain insufficiently accommodated.

Analysis of the neurodiversity-equality relationship must also consider additional underlying lines of reasoning. From a constitutional perspective, one assumes that individuals are born equals. However, by nature, human beings are inherently different. Some of these differences are irrelevant for people’s lives. Some impact individuals’ lives only because of bigotry, such as in the case of racism based on skin phenotypes, misogyny or sexual orientation. Others, such as disabilities, profoundly impact individual existence. In the case of diverse neurodevelopmental profiles, these differences collide with the way behavioral standards and social and legal norms


279. Perju, supra note 277, at 316–35 (explaining the lack of progress).

280. With different combinations and range of constitutionalized grounds of discrimination, all world constitutions protect equality and fight discriminations. This is granted through equal protection clauses (formal equality), substantial equality (also defined as indirect or disparate impact as well as indirect or systemic discrimination), and the duty of accommodation and affirmative actions. See generally Sandra Fredman, Comparative Study of Anti-Discrimination and Equality Laws of the US, Canada, South Africa and India (2012). Although the implementation of the idea of equality is an endless legal, political and cultural process, modern constitutions embody provisions limiting states’ power to adopt laws discriminating classes of people on the basis of race, gender, religion, nationality, ethnicity, origins, opinions as well as state statutes preventing horizontal discriminations perpetrated by private actors. For further discussion, see Ronald Dworkin, Sovereign Virtue: The Theory and Practice of Equality (2002) (analyzing the theoretical implications of social unequal distribution); Antony Flew, Equality in Liberty and Justice (1989) (discussing the tension between equal outcomes and equal justice); Fredman, supra note 162; David B. Oppenheimer, Sources of United States Equality Law: The View from 10,000 Meters, 10 Anti-Discrimination L. Rev. 20 (2010) (analyzing the anti-discrimination Law in the U.S.); O’Regan & Friedman, supra note 161.
are designed—the so-called neurotypical environment.\textsuperscript{281} In this context, it is unlikely that state legislations would deliberately discriminate against neurodiverse people. As a result, direct discrimination and violation of equal protection do not seem to be an actual threat, while indirect discrimination appears to be a much more problematic issue. Where neurodiverse traits are concerned, individuals suffer and are often excluded from social and civil citizenship. Substantial equality, through the expansion of education, health care and social services, should be consequentially strengthened.\textsuperscript{282} Providing services meant to allow neurodiverse people to expand their capacity to operate in social environments is, however, only a one-way solution oriented toward individuals.

A neurodiversity approach goes beyond this strategy. More accurately—by considering the \textit{adaptive} or \textit{maladaptive} nature of human attributes—it suggests operating in a bi-directional way, both toward individuals and social structures.\textsuperscript{283} While some NDD cognitive phenotypes are \textit{maladaptive per se} and should be accommodated with the \textit{medical model of disability}, other phenotypes associated with neurodiverse conditions are maladaptive due to social environments structures—the \textit{social model of disability}.\textsuperscript{284} Oliver Sachs saw this reality at work.

\begin{itemize}
\item \textsuperscript{281} Either atypical attributes are underestimated (or not considered) so individuals are equal no matter what diversities they display (anti-classification), or abnormal attributes determine that an individual is different and he should be treated differently (anti-subordination), triggering the risks of segregation and indirect exclusion. See Dana Lee Baker, \textit{Neurodiversity, Neurological Disability and the Public Sector: Notes on the Autism Spectrum}, 21 \textit{Disability & Soc'y} 15, 16 (2006).
\item \textsuperscript{283} Physical, biological, behavioral or neurocognitive. See Crespi, \textit{supra} note 12, at 6 (describing various phenotypes of autism).
\item \textsuperscript{284} See \textit{Baker}, \textit{supra} note 3, at 7 ("Conception of disability is socially relative and, therefore, at least somewhat unique to time and place . . . .
Observing the La Crete community in the United States, where there was a high prevalence of people affected by Tourette Syndrome, Sachs affirms that: "[t]he La Crete visit was extraordinary in many ways and broadened my sense of the range of Tourette’s syndrome and of people’s reactions to it. It also gave me a sense of how strongly Tourette’s, though neurological in origin, could be modified by context and culture . . . ." 285

Similarly, neurodiversity de-emphasizes the binary judgment of normal-abnormal and shifts the focus on an analysis of the functional interaction between human traits and social environments. This idea, cyclically discussed in psychiatry as well as in disability studies, assumes that social conventions consider only limited ways of performing cognitive tasks typical. It implies that only selected ways of learning, reasoning, socializing, perceiving stimuli, and processing emotions give people competitive benefits in schools, jobs, and social marketplaces. As a result, certain expressions of intelligence are labeled and stigmatized, precluding people from having access to social capital. 286 Due to those mechanisms, potentially adaptive traits are misconceived, downgraded, or lost. 287

Thus, the neurodiversity paradigm proposes being more inclusive of potentially valuable cognitive phenotypes. In contemporary complex societies, a new Theory of Justice parameter should be considered: the way through which cultures and legal systems perceive and select valuable cognitive capacities. It could be defined as brain equality. Traits are the result of unpredictable combinations of weaknesses and strengths, compensation and expansion of brain networks, speed and slowness in stimuli processing, neurophysiology, traumas, life ex-


286. See BAKER, supra note 3, at 216 (“For much of human history, the prevailing social context of the politics of neurodiversity and neurological difference has been the unquestioned hegemony of neurotypicality. Minds falling outside the established norm were at best tolerated as charmingly quirky, but were generally identified only as potential threats to sane society and human decency.”).

periences, education, and so on. How these traits perform depend on their interaction with modifiable social environments. By changing these environments, certain traits may become more adaptive or vice versa.

In conclusion, the notion of equality rooted in neurodiversity turns into an idea of fair bilateral endeavor. People with different NDD profiles, with adequate support, accept what it takes to be able to operate in society. As scientific research, clinical practice and biographical accounts demonstrate, divergent neuro-conditions differ enormously from one another. Every individual and family has to overcome extremely different hurdles. It is a challenge to translate this complexity into law. However, plaintiffs' claims for more educational services prove the commitment and the desire to move toward society. On the basis of the same endeavor—by reconsidering bias and misconceptions—society can behave in a bidirectional dynamic of expanding inclusion. Interpreted as bilateral commitment to move toward each other, the concept of neurodiversity contributes to deepening the debate around one of the core missions of modern constitutionalism, that is, establishing a societal sense of we balancing uniformity and diversity. The lack of cultural, political, and legal recognition of the phenomenology of human intelligence pluralism is detrimental not only for individuals who experience rejection, but also for societies and communities that ultimately lose tremendous human potential.

288. See generally Sapolsky, supra note 53.
289. See Putnam, supra note 269 (discussing the value of social solidarity).
Neuroethics Roadmap

The ACD Working Group on BRAIN 2.0 Neuroethics Subgroup (BNS) (https://acd.od.nih.gov/working-groups/brain2.0-subgroup.html) was formed to develop a Neuroethics Roadmap for the NIH BRAIN Initiative (https://www.braininitiative.nih.gov/); review priority areas identified in the BRAIN 2025 (https://www.braininitiative.nih.gov/sites/default/files/pdfs/brain2025_508c.pdf) (pdf, 1.2 mb) Strategic Plan, incorporating updates from the broader BRAIN 2.0 Working Group; and characterize the neuroethical implications that may arise as BRAIN Initiative investments produce new tool/neurotechnologies, and/or those tools/neurotechnologies are applied to advancing the goals of the NIH BRAIN Initiative.

The BNS has conducted a portfolio review and held a public workshop on neuroethical issues posed by research through the BRAIN Initiative. The BNS would now appreciate comments from the public on draft findings and analysis detailed in a Neuroethics Roadmap. The Neuroethics Subgroup has also provided analysis and findings to the Working Group on BRAIN 2.0 for inclusion in the Working Group’s initial thoughts. The public comment period on the Neuroethics Roadmap has closed. The findings and analysis from the BNS will be presented to the Advisory Committee to the Director, NIH, for consideration at a public meeting on Jun 13-14, 2019.

THE BRAIN INITIATIVE AND NEUROETHICS: ENABLING AND ENHANCING NEUROSCIENCE ADVANCES FOR SOCIETY

CHAPTER 1. NEUROETHICS PAST, PRESENT, AND FUTURE

CHAPTER 2. STUDYING OURSELVES: THE UNIQUENESS OF NEUROSCIENCE

CHAPTER 3. NEUROETHICAL ISSUES AND NEUROTECHNOLOGIES

The BRAIN Initiative is dedicated to revolutionizing the world’s understanding of the human brain through the development of tools, methods, and knowledge bases that will advance fundamental understanding of brain function – with a particular focus on circuit-level analyses. This work is
expected to lay the groundwork for a dramatically enhanced understanding of ways in which the brain can be coopted by disease, as well as provide new frameworks for effective interventions and therapies to treat brain disorders. However, as new neurotechnologies are developed, and new insights into the mechanisms of brain function and disease are discovered and refined, there is an obligation – and an opportunity – to continually consider, anticipate, and address potential neuroethical issues that may arise. In this way, neuroethics may be used not to impede, but to advance the ability of BRAIN research to have the greatest societal impact.

The first 5 years of the BRAIN Initiative saw progress in each of its designated research Priority Areas, some resulting in exceptional and unexpectedly rapid knowledge growth. BRAIN 2.0 will likely see significant advances in integrative strategies cross-cutting these Priority Areas, building upon groundwork laid during BRAIN 1.0.

**Table 6. Neuroethics Questions for Neuroscientists (NeQN)**

1. What is the potential impact of a model or neuroscientific account of disease on individuals, communities, and society?

1a. Possible unintended consequences on social stigma and self-stigma

1b. Possible social or cultural biases in research design or interpretation of scientific results?

2. What are the ethical standards of biological material and data collection and how do local standards compare to those of global collaborators?

2a. Protecting the privacy of human brain data (e.g. Images, neural recordings, etc.) and data, in immediate or legacy use beyond the experiment?

2b. Special regard for brain tissue and its donors due to tissue origin and its past

3. What is the moral significance of neural systems that are under development in neuroscience research laboratories?

3a. What requisite or minimum features of engineered neural circuitry generate concern about moral significance?

3b. Are ethical standards for research adequate and appropriate for evolving methodologies and brain models?

4. How could brain interventions impact or reduce autonomy?

4a. Identifying measures to ensure optimal autonomy and agency for participants/users

4b. Responsibility for effects (where responsibility broadly encompasses legal, economic, and social contexts)

5. In which contexts might a neuroscientific technology/innovation be used or deployed?

5a. Identifying applications that might be considered misuse or best uses beyond the laboratory?
5b. Does this research raise different and unique equity concerns and, if so, have equitable access and stakeholder benefit been considered?

In this chapter of the Neuroethics Roadmap, we employ the structure of these scientific Priority Areas for BRAIN as a framework for identifying companion neuroethical issues and neuroethics research opportunities. Identifying neuroethics issues is often not an intuitive process for neuroscientists. In anticipation of such issues it is useful to align scientific progress of BRAIN with the previously derived Neuroethics Questions for Neuroscientists (https://www.ncbi.nlm.nih.gov/pubmed/30308169) (NeQNs), developed by consensus at the Global Neuroethics Summit (see Table 6) in collaboration with many large-scale brain research efforts including members of BRAIN. As described in Section 1 of the Roadmap, these NeQNs can help focus attention on potential neuroethical issues and research opportunities that can then be judiciously addressed. To illustrate this rubric for the identifying neuroethical issues, the anticipated scientific advances highlighted in this section of the Roadmap will be cross-referenced with NeQNs that may be useful in eliciting any associated neuroethical concerns. As well, the frontier nature of neuroscience research also presents challenges related to unintended consequences deriving from their novelty including for example appropriate informed-consent procedures when it is impossible to quantify unintended consequences of controlling brain circuits, it is important to view neuroethical issues in context of the Neuroethics Guiding Principles (http://www.jneurosci.org/content/38/50/10586) (see Table 7) which can also be informed by the NeQNs. The principles offer a framework for prioritizing values and for ethical guidance for the conduct of BRAIN research including new technology development.

### Table 7: Neuroethics Guiding Principles

1. Make assessing safety paramount
2. Anticipate special issues related to capacity, autonomy, and agency
3. Protect the privacy and confidentiality of neural data
4. Attend to possible malign uses of neuroscience tools and neurotechnologies
5. Use caution when moving neuroscience tools and neurotechnologies into medical or non-medical uses
6. Identify and address specific concerns of the public about the brain
7. Encourage public education and dialogue
8. Behave justly and share the benefits of neuroscience research and resulting technologies

### Priority Area 1. Discovering Diversity

*What do we know?*
This BRAIN 2025 goal aims to develop a systematic and detailed understanding of the genetic, morphological, and physiological characteristics of different cell types throughout the nervous system as well as their potential roles in brain processes. Achieving this goal will permit development and use of genetic and molecular tools to identify cells – and ultimately, modulate their behavior in specific brain areas and circuits. This aspect of BRAIN 1.0 has been very successful, greatly expanding our knowledge of the number and diversity of cell types in the brain of different organisms, while also enabling us to quantify differences and likenesses between organisms.

What could we learn? Neuroethics research considerations

The rapid growth of technologies for gene editing (e.g., CRISPR), creating better viral vectors, tissue processing, imaging, and in-situ analysis of cells suggest that researchers are on the cusp of identifying and selectively modifying specific cell types, genes, or proteins in living systems. As research using models of human neural circuitry becomes more sophisticated, questions will arise about the appropriate boundaries for cell-type based manipulations that involve nonhuman animals (https://www.ncbi.nlm.nih.gov/pubmed/29203824), in particular (cite NHP genetic manipulation work) (see Chapter 4. Neuroethics and Research with Animals) as well as how increasingly sophisticated engineered neural circuitry and systems may challenge how we morally consider them (NeQN3).

Neuroethics Research Opportunity

What are the requisite or minimum features of engineered neural circuitry required to generate a concern about moral significance? (NeQN3a)

Collaborative research involving scientists, philosophers, and ethicists can define, create systematic approaches in conceptual and empirical work, and analyze testable measures of neurally derived features that would cause tension for donors, scientists, and members of the public. Importantly, exploring these ethical perspectives from a variety of stakeholders should incorporate multicultural research design and dissemination of information. The work should also explore the global relevance of these neuroethical considerations.

Research samples

Aside from general issues noted above, the origin of samples obtained for BRAIN Initiative-funded work invokes questions about the nature of the samples, applicability and privacy. Related to the former, initial work on discovering and characterizing cell types should not focus on a single species, or one societal group or sex or race or age, ensuring that bias is mitigated, and that the
benefits of neuroscience research can apply to individuals from numerous populations (NeQN1b).
Related to privacy, consent from individuals who provide tissues, either while alive or after death,
should address the long-term consequences of such a donation for themselves and their relatives.
These issues should be addressed within the informed-consent process and on a continuing basis
should unintended information be derivable from these tissues such as stigma or broader
application (NeQN1b, NeQN5b, Guiding Principles 3, 8). Important to note that in an era of data
sharing, it may be the legacy use of the data derived from the sample, rather than just the tissue
itself that warrants privacy consideration (NeQN2a). Further, as research models evolve, they may
warrant additional scientific and ethical review: At what point should in-vitro or ex-vivo human cells
or samples be considered to warrant greater moral significance or revised research standards?
These issues raise neuroethical questions about moral significance that should not only involve
scientists and ethicists, but also incorporate how to address concerns the general public may have
(NeQN3a,b, Guiding Principle 6).

Cell/tissue manipulation

This cell-census component of the BRAIN Initiative raises familiar ethical, legal, and social
implications that have already emerged from genomic research; for example the ability to introduce
whole genes into cells warrants forethought about the resulting effects on function (e.g., on circuits,
both in the recipient individual but also in later generations) (NeQN5a, Guiding Principle 1).
However, there may be unique considerations with manipulations that alter brain function. Along
with decisions on which cells should be targeted (healthy or diseased), a framework will be needed
to revisit aspirations for these methodologies and what distinguishes restoring cognitive health
versus enhancing cognitive function or altering learning and memory (NeQN1a, NeQN5, see also
CHAPTER 5: BEYOND THE BENCH: REAL-WORLD TRANSLATION OF NEUROSCIENCE
RESEARCH). The ability to manipulate cells and tissues may enable researchers to understand
more about cell identity – and in time, perhaps individual identity. How does this bear on privacy for
not only individuals who participate in studies, but family members as well? (NeQN2a, Guiding
Principle 3).

Priority Area 2. Maps at Multiple Scales

What do we know?

This BRAIN 2025 research area is focused on developing detailed knowledge of the structural and
functional spatial representations of activity and interconnectivity within the brains of different
model organisms, at scales ranging from individual synapses to large-scale connectivity of human
brain regions. This project has supported the development of significantly enhanced methods for
generating structural and functional maps from ex-vivo and living brains in species ranging from
worms and flies to humans. Although improvements to non-invasive human-brain imaging
technologies during BRAIN 1.0 have been incremental, technologies such as fMRI are improving in
speed and signal-to-noise ratio. Moreover, portable, near-infrared spectroscopy can already
provide non-invasive readouts of brain activity in social settings (and portable PET and MRI
systems are in development). These enhanced imaging methods, combined with studies across species, could soon reveal functional activity and connectivity patterns that may potentially be interpreted in terms of the capacity for thought, mood states, behavior, and personality – directly from physical observations of the human brain. As described in the BRAIN 2025 report, an eventual goal is to discover how the human brain produces cognition and behavior at the “speed of thought,” information that could inform how people make decisions that form the basis of personality and self. The wider availability of such techniques for human use may prompt non-medical, commercial, consumer, or judicial use of such technologies and care will be needed to determine which contexts technology or innovation can be justly deployed (NeQN5, Guiding Principle 8; see Chapter 5: Beyond the Bench: Real-World Translation of Neuroscience Research).

What could we learn? Neuroethics research considerations

In this Priority Area, similar neuroethical considerations apply to those detailed in the Discovering Diversity Priority Area related to sample selection and use, along with the need to protect information that could plausibly be extracted from tissues donated by human research participants or patients. Distinct in this Priority Area are neuroethical questions related to mapping studies. For example, currently, transcranial direct current stimulation is being used in non-medical, non-research settings toward improving neurological performance. Such issues are not directly within the scope of current BRAIN Initiative-funded research, but they are relevant to consider as federally funded research finds application in everyday life (see Chapter 5: Beyond the Bench: Real-World Translation of Neuroscience Research, NeQN5, Guiding Principle 5).

A pervasive challenge with all research involving humans is defining “normal” in the context of health and disease, but also in the context of human variation and individual identity and personality. Scientific studies in nonhuman and human animals using male-only samples have sometimes generated incorrect general assumptions resulting in adverse health consequences (https://www.sciencedirect.com/science/article/pii/S0039606014004255) or reinforced negative biases about socially constructed groups already present in society. These factors should be taken into consideration to ensure that large-scale studies to understand the brain’s maps and networks both sample from (and can thus benefit) a fully representative cross section of our society. Communication and use of the results of such studies should also be carefully managed to ensure that the design of experiments and interpretation cannot be subverted to fuel existing negative societal biases or prejudices (NeQN1a,b, Guiding Principle 7).

Neuroethics Research Opportunity

How can human brain data (e.g., images, neural recordings, etc.), and the privacy of participants from whom data is acquired, be protected in case of immediate or legacy use beyond the experiment? (NeQN2a)
There is an opportunity for collaborative study about the scientific capabilities of BRAIN Initiative research as well as consideration of legal definitions and historical and evolving public views about neuroprivacy. Part of the assessment of public views could involve exploring new types of informed-consent processes, in particular, for research involving neural recordings. Other projects could explore and assess best practices for community engagement and communication strategies with neuroethical issues on stigma, bias, and privacy. Cultural views across and within cultures and geographic regions will provide greater insight into such technologies might be received and used within a global landscape.

**Priority Area 3. Brain in Action**

*What do we know?*

This BRAIN 2025 research area aims to identify and understand neural activity patterns that underlie cognitive processing and behavior. The BRAIN Initiative has supported many recent advances enabling recording and modulation technologies that are to be used in non-human animals. Studies are now deploying new technologies for large-scale recording of multiple variables (including neural activity and neurotransmitter concentrations) within non-human animals engaged in complex tasks or more naturalistic behaviors compared to previous anesthetized or head-fixed activities. A parallel area of significant growth has been the application of machine vision and deep-learning approaches to large-scale quantification of non-human animal behavior – from counting the number of times a fruit fly grooms itself in a dish to tracking the paws and whisker movements of a mouse during real-time brain imaging. While much of this work and preliminary insights are from nonhuman animals, the aspiration is to expand these abilities, with greater precision and to larger numbers of recorded neurons, including novel non-invasive imaging technologies, into humans.

*What could we learn? Neuroethics research considerations*

New combinations of brain activity and behavioral data are beginning to enable development of models and theories that can depict and reproduce the brain’s computational codes that lead to complex behaviors. In this Priority Area, similar neuroethical considerations about potential for bias and stigma as well as neuroprivacy apply as detailed in Priority Area 2. Maps at Multiple Scales. One distinguishing feature is the issues around developing non-invasive brain recording devices.

Wearable human-brain imaging technologies are permitting brain to behavior correlative studies in humans. Knowledge emerging from these studies could allow us to assess mood states, behavior, and personality directly from physical observations of the brain, or even from assessment of physical behavior in different environments. Therefore, the same careful considerations will be needed related to risks of reinforcing bias by dividing participants along socially constructed identities as well as with privacy (NeQN1b and NeQN2a). Wearable neurotechnology is not only an
interest of biomedical researchers, but it is already an area of active exploration in the commercial sector as a wellness or cognitive enhancement tool. While not the intended context for BRAIN research, insights from BRAIN will likely extend beyond the BRAIN community and its mandate. Ethical stewardship of studies exploring “brain in action” maps of designated ‘normal’ and ‘abnormal’ brains, particularly as they relate to mental health as well as implications for enhancement will require considerations of privacy and best uses and possible restricted uses beyond the biomedical setting (NeQN5 and Guiding Principle 5). To be clear, the weight and responsibility of this is not work for the neuroscientists alone. These questions of uses “beyond the bench” are best explored as a multi-stakeholder project. (see box). In addition, performing best ethical practices also relies on mechanisms and infrastructure to support the scientists’ ability to do so.

**Neuroethics Research Opportunity**

In which contexts might a neuroscientific technology/innovation be used or deployed? (NeQN5)

Ethical stewardship of neuroscience and its products requires a scientist’s involvement in anticipating best uses and possible misuse. However, to identify frameworks for best use and misuse, research should involve an anticipatory approach scanning the horizon for possible contexts for use in the near and intermediate future as well as exploring existing ethical and legal guidelines with a diverse set of stakeholders including end users, consumers, scientists, ethicists, legal scholars, as well as members from the policy community. Importantly, the exploration of this anticipatory work should include a global community who may have differing values and priorities for the use of such research findings and developments that may be at odds with national views and policies.

**Priority Area 4. Demonstrating Causality**

*What do we know?*

This BRAIN 2025 research area aims to test our understanding of the brain through perturbations that lead to predictable outcomes (cause and effect). This research encompasses technological development and refinement of experimental methods such as optogenetics that permit specific cells within the brain to be turned on or turned off or chemogenetics that enables pharmacological manipulation of specific cells—enabling evaluation of the immediate and long-term effects of these perturbations on brain function or behavior. A direct analogy in humans is deep-brain stimulation used to alleviate Parkinson’s tremors. As techniques for performing manipulations of brain-cell activity become more refined, selective, and deployable, our understanding of how such manipulations affect brain function and behavior is also becoming more sophisticated. At the heart
of this priority area is the development of interventional technology that can manipulate the brain for desired behaviors with an ultimate goal that insights will provide ways to relieve unwanted brain function and behaviors that arise from brain diseases such as mental illness or degenerative disease. In order to ensure that such insights and abilities to intervene with the brain result have the greatest impact for alleviated suffering, careful consideration should be made of how such interventions may intentionally or unintentionally impact or reduce autonomy, capacity, and agency (NeQN4, Guiding Principle 2).

What could we learn? Neuroethics research considerations

Unknown consequences of manipulation.

Although physiological manipulation of cells and tissues is currently a valuable tool for neuroscientists seeking to understand brain circuits in non-human animals, and in some cases in humans, these techniques hold significant potential for therapeutic use. For example, they could also be used to intentionally augment, restore, and/or redirect brain function (see Priority Area 6, below). Current genetically targeted methods of brain-cell manipulation such as optogenetics require genetic modification of a non-human animal through breeding or viral transfection, as well as delivery of intense light to the location requiring “activation.” Other technologies such as chemogenetics also require genetic modification of specific cells but then only require systemic drug delivery for activation. While most researchers do not envision this technology being used in humans, similar genetic therapies are being explored (including optogenetics in the eye as a treatment for blindness). What is more likely in humans are invasive studies including deep brain stimulation, multi-electrode array recordings from the surface of the brain, and noninvasive stimulation such as ultrasound and transcranial magnetic stimulation. With any of these technologies, the aim would be to manipulate or control the brain in a way that a patient or participant could not do on their own. In other words, there’s a challenge to the user’s autonomy and agency. Therefore, at times, there may be a need to explore how measures can be put in place to ensure an optimal level of autonomy and agency for users (NeQN4, Guiding Principle 2). Part of this discussion should involve exploration of how measures for override or user control of stimulation parameters can be offered to users that would be not only beneficial and desired by participants, but also balanced with what scientists understand are optimized parameters for use in treating particular aspects of disease. Further, considerations for safeguards from hacking or misuse as well as understanding of who takes ultimate responsibility for ongoing support for the technology beyond the lifetime of a research project and unintended consequences of device use will also maximize technologies developed under this priority (NeQN4b, Guiding Principle 4).

Neuroethics Research Opportunity
Debate about cognitive enhancement has been active for many years. See, for example, *The President’s Council on Bioethics. Beyond therapy: Biotechnology and the Pursuit of Happiness* ([https://biotech.law.lsu.edu/research/pbc/reports/beyondtherapy/](https://biotech.law.lsu.edu/research/pbc/reports/beyondtherapy/)). Deeper exploration is warranted about the scientific possibilities and limits of today’s and tomorrow’s neuroscientific advances – and the conceptual separation between therapy and enhancement. Collaborative research involving scientists, ethicists, legal scholars, and practitioners exploring evolving societal definitions of disease and aspirations for wellness as well as research involving ethical and legal standards in this space on a global scale are needed. These conceptual ethics approaches can also be complemented by public engagement research exploring public awareness, opinions and assumptions about neuroscience and enhancement. One significant question to explore with regard to neurotech development in this space is: Does neuroscience raise different and unique equity concerns and, if so, have equitable access and benefit of stakeholders been considered? (NeQN5b)

*Risk analysis*

Research is needed to understand the unique health and safety risks, as well as potential unanticipated consequences of an intervention, on a person’s autonomy, capacity, and agency including those related to altering features of personality and memories (NeQN4a, Guiding Principles 1,2). In addition, a valuable part of this risk analysis might include comparisons of newly developed BRAIN intervention with existing ones and even re-evaluation of older ones based on new insights derived from research in BRAIN. A better understanding of how existing interventions, even psychostimulants, affect the brain, acutely and long-term, is sorely needed and should represent a backdrop for considering the advantages and disadvantages of new interventions. It may be that focal interventions are safer than pharmacological interventions in terms of side effects. Importantly, features of ‘risk’ should include not only physical harms, but also social ones. Given the potential complexity of the effects of interventions on autonomy, agency, and capacity, neuroethical research combined with scientific efforts could help sort out what participants and scientists understand about these terms and how they evaluate their importance in the context of these interventions. Because participants may have difficulty understanding the unique risks that result from manipulating circuit function, informed consent processes may warrant deeper review (NeQN4, Guiding Principle 2).

**Neuroethics Research Opportunity**

Can neurotechnology be designed with technological safeguards that enhance an individual’s autonomy, or that protect negative impacts on his or her agency? (NeQN4)
Such research should involve a mixed team of scientists, ethicists, and future end-users. Collaborative research evaluating current societal, ethical, legal meanings of responsibility when technologies function well (or when they do not) could help inform new best practices and guidelines for this type of research. Approaches would include conceptual work on understanding divergent and shared meanings for these terms for stakeholders as well as empirical work assessing understanding, values, and preferences in BRAIN-Initiative funded research.

**Priority Area 5. Identifying Fundamental Principles**

The BRAIN 2025 report identified a central role for data analysis, theory, and modeling for the purpose of extracting information from data sets, and for developing conceptual and algorithmic frameworks for interpreting circuit dynamics underlying key brain processes like sensory processing, motor control, and decision-making. As noted above, such work will ultimately provide the conceptual backbone for interpretation of data and ultimately understanding of how the brain functions and malfunctions. Neuroethical issues and research in this Priority Area overlap with those described in the other Priority Areas, above, and is not repeated here. However, data sharing is a key element associated with this Priority area, as elaborated further below, sharing raises ethical concerns familiar to any field which collects large datasets, but also could raise a greater degree of tension due to the potential sensitivity of brain-based data.

*Data sharing*

Large amounts of data are required to enable development and testing of theories and models. However, the imperative for experimentalists to share their data – and the need for others to mine and extract information from highly complex, multi-dimensional and multi-modal data – presents significant challenges and has a topic of significant discussion among scientists. Data sharing – including analyses, algorithms, and shared access to infrastructure – is an essential component of open and equitable science (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3102049/). It is also a hallmark of rigorous and ethical research. While many of the challenges related to data sharing are technical, just as significant are the social and cultural challenges, i.e. convincing a profession that relies on the currency and attribution of peer-reviewed publications to openly share data without the guarantee of “credit” raises tensions in any discipline. Therefore, there are legitimate concerns of scientists that must be addressed in addition to the need to provide support and incentives to engage in data sharing.

Neurotechnologies continue to become more sensitive, more robust, more portable, and multimodal. The proliferation of neuroscience into society offers great promise for new insights and improved social policy. But with these dramatic and rapid shifts come difficult ethical questions...
about the collection, interpretation, application and access of scientific data. A key component to this Priority Area is the aggregation of data collected across large numbers of non-human animals, humans, labs, and institutions.

NIH has several data-sharing policies, as does the BRAIN Initiative itself (https://grants.nih.gov/grants/guide/notice-files/NOT-MH-19-010.html). These policies have been implemented in multiple ways, including developing a central repository for data, standardized analysis procedures, and policies for data sharing. Responsible data sharing promotes equity, whereas exclusion of data can lead to knowledge confined to a limited group of individuals. Purposeful exclusion of data may lead to hypotheses that reinforce previously held biases (NeQN1b).

Brain data come in many formats, including measurements from genetic, genomic, protein, functional, imaging, and behavioral analyses. These different data types are usually compiled into specific databases with specific standards. As the need for integration across data types and platforms evolves – a key goal for BRAIN 2.0 – it may be necessary to revise data-sharing policies to encompass the widening utility of the data. For human data, participant-privacy issues have always been paramount, primarily with regard to identity. We consider human brain data different because of its potential to gain insight into an individual's thoughts and other aspects of an individual. As the BRAIN Initiative moves ahead (and neuroscientists work across the globe outside of the BRAIN Initiative), increasing amounts of data will accumulate from diverse experimental approaches that will likely be more individually precise – and potentially more identifiable. We need to understand more about public awareness and concern about brain privacy (see Chapter 5: Beyond the Bench: Real-World Translation of Neuroscience Research, NeQN2a).

Should all brain data be shared?

When neuroscience data are used to investigate brain function, ethical use of the data requires i) noting its source; ii) insuring that it was properly obtained according to ethical guidelines and university, company and/or local, national, or international statutes; iii) using only the subset of data required to query the question of interest; and iv) properly acknowledging the data source. However, there are circumstances that may preclude universal data sharing. These special circumstances include, for example, when a research participant's identity could be compromised from combining that individual’s composite datasets, which was neither envisioned nor specified in the informed-consent process.

Neuroethics Research Opportunity

What are the ethical standards of biological material and data collection and how do local standards compare to those of global collaborators? (NeQN2)

The Human Brain Project has produced a report on data sharing, privacy, and practices moving forward as informed by exploration of conceptual analyses of privacy, public opinions on privacy, as well as technical and legal analysis. Not only does this represent
the type of interdisciplinary work that BRAIN’s neuroethics research should strive to do, they represent a rich opportunity for collaboration. Their current activities related to the ethics of data sharing align with similar goals of the BRAIN Initiative and could provide a fertile ground for research on best practices for neuroethics research, neuroscience data collection, and public engagement. International collaboration and active dialogue about these practices will also be critical as much of these data sharing practices occur on a global backdrop and will require deeper reflection of the ethical standards of data collection nationally and how these compare to those of global collaborators (NeQN2).

Big data practices as well as data access have transformed the possibilities of the applications of data and what kinds of information can be derived. The purpose of big data analytics is to create new unanticipated knowledge and in so doing, appreciating what information can be derived from data and anticipating risks (https://journals.sagepub.com/doi/full/10.1177/2053951716650211) is a significant challenge. One way to address this issue is to have routine review in collaboration with scientists, data analysts and ethicists to evaluate how new analyses might open new opportunities for risk, particularly of re-identification.

While IRB approval is an important practice for conducting research with human participants, IRBs are not uniform in considering the special circumstances associated with BRAIN Initiative-derived data. In cases where collecting brain data is the goal of an experimental protocol, local IRBs should obtain neuroethics input, thereby ensuring that the consequent data have been vetted for these concerns. As the data become more complex as a result of combining different data types, more information about human participants including unintended data disclosure will be decodable. Anticipating the impact of the availability of these data is difficult and will likely pose new neuroethical concerns. Given this changing data and analysis landscape, it would be prudent for institutional and Office for Human Research Protections (OHRP) IRB guidance to be revisited on an ongoing basis to insure that current guidance sufficiently protects the participant, and if not, then new guidance should be proffered to deal with the associated issues.

International collaboration and active dialogue about these practices will also be critical as much of these data sharing practices occur on a global backdrop and will require deeper reflection of the ethical standards of data collection nationally and how these compare to those of global collaborators (NeQN2).

**Neuroethics Research Opportunity**

What are the possible unintended consequences of neuroscience research on social stigma and self-stigma? Is it possible that social or cultural bias has been introduced in research design or in the interpretation of scientific results? (NeQN1)
Develop practices to enhance inclusiveness and reduce bias. Studies should be designed to investigate the impact of many variables on brain function, including but not limited to, sex, race, and cultural experiences. Explicit attention should be given to questions about who will benefit from neuroscience research advances, and how to promote equitability across these and other important domains. Neuroethical deliberation is necessary and requires thoughtful input beyond neuroethics alone – including, for example, experts in sex/gender differences, cultural and societal differences, disease advocacy, and other topics related to human variation.

Priority Area 6. Human Neuroscience

This BRAIN 2025 research area aims to develop innovative technologies to understand the human brain, with the ultimate goal of treating its disorders. In this Priority Area, new technological and conceptual approaches are integrated and applied to discover how dynamic patterns of neural activity become cognition, emotion, perception, and action in both health and disease. Such neurotechnologies can be used to monitor the brain to understand details of how it works in health and disease, as well as to design neurotechnologies to treat brain dysfunction. Currently, there is nothing that replaces the human brain as a model for understanding high-level complex outputs of the brain such as cognition. While human studies are conducted judiciously, studies in humans are often considered the most ethically complex (see Chapter 2. Studying Ourselves: The Uniqueness of Neuroscience).

What could we learn? Neuroethics research considerations

Human samples and recordings

Most human brain samples (and invasive recordings) come from diseased brains in which intervention has been warranted, or after the individual has died. Neuroethical research should examine how to define acceptable ways to acquire tissues and measurements from healthy brain tissue that can provide valuable and necessary control information (NeQN1a). Additional studies can explore how to coordinate efforts between researchers (including internationally) to ensure that large-scale census type work spans sufficiently diverse populations to ensure equitable benefit (NeQN1b). Other topics for consideration include the potential need for guidelines pertaining to use of human cells in the context of multicellular assemblies such as organoids and assembloids that, as the complexity increases, may attain moral status (NeQN3 as mentioned in Priority 1 and Chapter 2: Studying Ourselves: The Uniqueness of Neuroscience).

Noninvasive recording and imaging

The ability to perform non-invasive neuroimaging presents a number of areas of potential neuroethical concern, including: i) unexpected access to incidental findings; ii) detection of clinical biomarkers of latent or impending disease; and iv) use of neuroimaging for national security, legal, and marketing activities. Imaging for clinical biomarkers is becoming more common, and it will
likely be combined with self-reports and expert observations to better evaluate clinical state. Neuroethical research could explore the ethical consequences when neuroimaging results diverge from what an individual research participant or patient experiences and what his or her provider gleans from the technology. Other questions surround decision-making related to data sharing and how that may differ among disease contexts (e.g., concussion vs. depression, NeQN1a). Another scenario to consider might be the prospect of identifying the potential to develop disease – such as Alzheimer’s – decades before symptoms appear. This may have implications for employment and insurance coverage, for example (NeQN1a).

Both implanted electrodes and noninvasive approaches that generate behavioral and neural recordings may uncover decodable information that can create privacy violations for an individual and at times family members as well. As discussed in Priority Area 3, these potential risks raise questions about appropriate use, sharing, and protections for data beyond its first use in an experiment. Importantly participants must have the knowledge and realistic expectations for de-identification in order to consent to providing their data. This is particularly important as technology and decoding algorithms as well as coupling with other experimental modalities may advance to the point that it may be possible to interpret brain states without a research participant’s permission (NeQN2a and discussed above in Priority Area 3). Importantly, algorithms, which may be assumed to be objective, should be explored and acknowledged for the possibility of carrying the inherent and unrecognized biases of their creators (NeQN1b).

Noninvasive neuromodulation

Invasive human recording and modulation

Brain-computer interfaces are already in widespread use; the work of the BRAIN Initiative will accelerate their development and the precision with which they can influence brain function. One current example is use of deep-brain stimulation to shorten or block seizures. New technologies under development will likely have the capability to monitor neural or neurotransmitter activity over long periods of time and to provide detailed patterned stimulation in a closed-loop (operator-independent) manner. This means that monitoring and manipulation occur in real time without control of the individual wearing the device.

Devices that are implanted into the brain entail a high level of risk, as they inevitably create an intimate connection between a device and an individual – along with risk of infection, rejection and the need for long-term care of the recipient and maintenance of the device. This reality raises significant and immediate neuroethical questions. The detail with which brain states can be monitored will likely improve as we obtain more robust/sensitive recording technology, such as flexible mesh electrodes that detect and potentially modulate electrical activity of many cells and can thus displace current limitations of available electrodes, such as static immobility and stiffness. The impact of these technologies is likely to increase dramatically with nanotechnologies that bring innovation related to materials science, optics, chemistry, and learning algorithms.
It is of note that the health and well-being of participants in invasive technology research presents heightened neuroethical concerns. If a participant benefits from the technology, then the question of what happens at the end of the study becomes more salient. Does the device remain implanted and functioning? Or what happens if it is removed, which would likely return the participant to his/her pre-therapeutic status? If the device remains who is responsible for its maintenance, ensuring that it functions well and monitoring of the participants health consequences long-term? These are important questions as the participant may be particularly vulnerable at the time of informed consent. Further there are long-term consequences beyond the participant including those for his/her immediate family.

**Neuroethics Research Opportunity**

Collaborative research projects between neuroscientists and neuroethicists could explore how to define and operationalize in a lab setting terms and features of sentience. These studies could also work toward developing technologies to measure sentience and other features. See also Neuroethics Moonshot.

Neuroethical research is needed to address several scenarios that will likely arise sooner than later. Examples include the value and risks of medically unnecessary work such as implantation of an experimental device that alters brain activity in healthy individuals and informed consent processes for neurosurgical patients for research associated with, but not necessary for medical care. In addition, broader questions that warrant deeper exploration are the long-term responsibilities for scientists, funding agencies, or device companies who implant devices in research participants as well as ownership and rights of participants to access to data from an implanted device.

*Conclusion*

The examples provided above illustrate neuroethical considerations framed by opportunities to integrate neuroscience research with what we expect to learn (as well as where it is difficult to anticipate what to expect) about understanding the human brain and behavior. When considering the important issues of agency, self, emotions, decision making – and even more familiar issues of learning and memory and consciousness – it is important to recognize that the biological underpinnings of these aspects of our personhood remain obscure (see Neuroethics Moonshot). Yet, there is a moral imperative to use the knowledge gained from the BRAIN Initiative to alleviate suffering from brain diseases and disorders. Intellectual freedom for scientists must be coupled with individual and institutional responsibility to assure morally responsible behavior – as well as establishing practical and sensible ways to assess societal benefits, safety, and security risks both before and after research, and limiting scientific projects and exploration when necessary. The
neuroethical impacts and implications of BRAIN Initiative-funded research should be assessed on an ongoing basis. Further, ongoing efforts contributed by the global neuroethics community to define appropriate limits and develop concrete guidelines are essential.
Interventional Tools (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=3)

Theory & Data Analysis Tools (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=4)

Human Neuroscience (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=5)

Integrated Approaches (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=19)

BRAIN Update (https://www.brainupdate.nih.gov/)

BRAIN 2025 Report (/strategic-planning/brain-2025-report)
NIH Institutes & Centers

NCCIH (https://nccih.nih.gov/)  
NEI (https://www.nei.nih.gov)  
NIA (https://www.nia.nih.gov)  
NIAAA (https://www.niaaa.nih.gov/)  
NIBIB (https://www.nibib.nih.gov)

NICHD (https://www.nichd.nih.gov)  
NIDA (https://www.drugabuse.gov)  
NIDCD (https://www.nidcd.nih.gov)  
NINDS (https://www.ninds.nih.gov)  
NIMH (https://www.nimh.nih.gov)

BRAIN Initiative Partners

Alliance Members and Affiliates (http://www.braininitiative.org/alliance/)  
All Participants (http://www.braininitiative.org/participants/)

More Information

Email Us (/about/contact-us)  
Archives (/archives)  
BRAIN Update Blog (https://brainupdate.nih.gov/)

Related Websites

National Institutes of Health (https://www.nih.gov)  
USA.gov (https://www.usa.gov)

Other Links

Adobe Acrobat (https://get.adobe.com/reader/)
Neuroethics Roadmap

The ACD Working Group on BRAIN 2.0 Neuroethics Subgroup (BNS) (https://acd.od.nih.gov/working-groups/brain2.0-subgroup.html) was formed to develop a Neuroethics Roadmap for the NIH BRAIN Initiative (https://www.braininitiative.nih.gov/); review priority areas identified in the BRAIN 2025 (https://www.braininitiative.nih.gov/sites/default/files/pdfs/brain2025_508c.pdf) (pdf, 1.2 mb) Strategic Plan, incorporating updates from the broader BRAIN 2.0 Working Group; and characterize the neuroethical implications that may arise as BRAIN Initiative investments produce new tool/neurotechnologies, and/or those tools/neurotechnologies are applied to advancing the goals of the NIH BRAIN Initiative.

The BNS has conducted a portfolio review and held a public workshop on neuroethical issues posed by research through the BRAIN Initiative. The BNS would now appreciate comments from the public on draft findings and analysis detailed in a Neuroethics Roadmap. The Neuroethics Subgroup has also provided analysis and findings to the Working Group on BRAIN 2.0 for inclusion in the Working Group’s initial thoughts. The public comment period on the Neuroethics Roadmap has closed. The findings and analysis from the BNS will be presented to the Advisory Committee to the Director, NIH, for consideration at a public meeting on Jun 13-14, 2019.

THE BRAIN INITIATIVE AND NEUROETHICS:
ENABLING AND ENHANCING NEUROSCIENCE ADVANCES FOR SOCIETY

CHAPTER 1. NEOUROETHICS PAST, PRESENT, AND FUTURE

CHAPTER 2. STUDYING OURSELVES: THE UNIQUENESS OF NEUROSCIENCE

CHAPTER 3. NEOUROETHICAL ISSUES AND NEUROTECHNOLOGIES

CHAPTER 4. NEOUROETHICS AND RESEARCH WITH ANIMALS
CHAPTER 5. BEYOND THE BENCH: REAL-WORLD TRANSLATION OF NEUROSCIENCE RESEARCH

CHAPTER 6. INTEGRATING NEUROETHICS AND NEUROSCIENCE

Neuroscience has captured the imagination of the non-scientific public and scientists alike because of the rich implications of its findings. As we have shown throughout this report, neuroethics is integral to the neuroscientific endeavor based upon unique ethical issues that arise in conjunction with assumptions and beliefs about the role of the mind and the brain’s connection to it. Neuroethics helps guide neuroscience advances and discoveries toward positive social outcomes – in medical or non-medical settings. In turn, the numerous impacts of neuroscience and neurotechnologies on individuals and populations have significant and broad-reaching ethical implications. In this chapter, we provide suggested concrete steps for integrating neuroethics into the study and practice of current and future neuroscience research.

Guidance for researchers: The critical role of neuroethics in neuroscience research

What do we know?

Neuroethics should be integrated into the entire life cycle of a neuroscience research project – from hypothesis to research design and conduct to dissemination of results and translation of knowledge. Many entities have long-recognized the importance of this interdependence, including at the highest levels of government. For more details, see Chapter 1. Neuroethics Past, Present, and Future.

While neuroethicists can work independently from neuroscientists, neuroethical expertise is most relevant with shared intimate knowledge of the science and its context. Including an ethicist in a research team can lead to fruitful inquiry and provide an opportunity for ethicists to not stall, but accelerate, good neuroscience by anticipating and addressing ethical issues before they arise. Such ongoing interactions mitigate potential roadblocks that ethical missteps create if not considered early and often.

What could we learn? Neuroethics research opportunities

Integrating collaborative neuroethicists within research teams has been the subject of several BRAIN Initiative-funded neuroethics research project grants (R01s). For example, one project explored the ethics of research involving brain organoids. Researchers also investigated the informed-consent process associated with invasive brain interventions (such as deep-brain stimulation) for psychiatric conditions or opportunistic research occurring with epileptic patients implanted with electrodes.

What is the best way for neuroethical questions to surface in neuroscientific settings? Neuroscientists need the ability to identify ethical quandaries in the context of their work. Although tides are changing, most neuroscientists do not know enough about neuroethics to navigate these waters. As a result of mandatory training, however, most neuroscientists are indeed aware of basic issues related to the responsible conduct of research. We have learned through such efforts to educate neuroscientists...
about neuroethics (particularly in preliminary conversations through the NIH BRAIN Neuroethics Working Group) that neuroscientists typically welcome a resource to help them explore the ethical, social, and legal implications that may arise uniquely because their subject of study is the brain. The Neuroethics Questions and the Neuroethics Guiding Principles (see Chapter 1. Neuroethics: Past, Present, and Future) help to serve this role. In addition, the international brain community has committed to addressing neuroethical issues in formal and informal ways. In a special issue on neuroethics in the journal *Neuron* (https://www.cell.com/neuron/fulltext/S0896-6273(19)30068-6), each of seven large-scale brain research projects demonstrates how they are currently or are planning to integrate neuroethics into their research projects.

**Neuroethics scholarship and training**

Do neuroscientists have enough training to understand ethical implications of their research? Are there established practices for professional conduct in highly innovative areas in which neurotechnological capabilities are surfacing rapidly?

*What do we know?*

Culture change is a key component to fully integrating neuroethics into neuroscientific practice. The next step will be to establish formal opportunities for established scientists and trainees conducting neuroscience research to learn about neuroethics, and for neuroethicists to learn more about neuroscience. Some of this training may fit well into core principles already articulated by the BRAIN 2025 report, stating the need for crossing boundaries to promulgate interdisciplinary research. To facilitate these interactions, the NIH Neuroethics Working Group has published a set of Guiding Principles as values to undergird neuroscience research (see Chapter 1. Neuroethics: Past, Present, and Future).

To motivate true scholarly partnerships, institutional support and incentives are needed— a structure, with resources supplied for both groups. Dedicated support may encourage such collaboration and give collaborating ethicists time to engage in co-developing neuroscience. These efforts could also help meet the BRAIN Initiative’s goal of breaking down “silos” between fields of study.

*What could we learn? Neuroethics research opportunities*  

Use of the Neuroethics Guiding Principles and the NeQN set provides both neuroscientists and ethicists a springboard to discuss the design, conduct, and translation of neuroscience research. For example, in NeQN1, scientists are prompted to consider how the questions they choose to study in the lab might amplify existing biases. Thus, considering these questions might lead them to reconsider designing a tool that uses a skewed mix of research participants (such as all males) as a normal population – a decision that may confound results. Interpreting such results will have implications for defining fundamental qualities of personhood associated with the brain, so choosing an appropriate study population is vital. Using NeQN4 and Guiding Principle 2, scientists are prompted to consider how brain interventions might affect autonomy. Researchers can respond to this question by designing technologies that enable ways an affected individual could override the machinery. Alternatively, researchers might also evaluate a neurotechnological design as impractical, by running down a battery, for example.
All true partnerships are two-way streets. While neuroscientists can benefit from learning more about ethics, ethicists must also keep apprised of current principles and trends in scientific and engineering research to have a better sense of how to navigate the ethical challenges. Having neuroethicists collaborating at the formative stages of BRAIN Initiative-funded research study design facilitates not only ethical neuroscience, but also provides more opportunities to speak a common language. Finding common ground will undoubtedly add fresh perspectives to conceptualizing, conducting, and translating research for the broadest number of people.

When should neuroethics education be integrated?

Among the neuroethics-related short-term goals outlined in BRAIN 2025 are to i) establish training grants for human research/ethics and ii) to establish neuroscience/ethics training programs, meetings, and interactions to establish guidelines and principles for human neuroscience research.

Ideally, as noted in Gray Matters Volume 1 (https://repository.library.georgetown.edu/bitstream/handle/10822/709231/Gray%20Matters%20Vol%201.pdf?sequence=1), exposure to ethics, and neuroethics, should happen early and often in a scientist’s professional development:

“Early ethics education in academic settings is critical to prepare future scientists to integrate ethical considerations into their work – including future research in neuroscience. Professional development for experienced investigators is equally important and can serve multiple ends, contributing not only to their individual knowledge, but to the knowledge of the students and young scientists that they mentor as well. Ethics education has a better chance of informing action when it is continually reinforced and connected to practical experience. (11, p.28)

and

“One foundational approach to integration is pairing science and ethics education at all levels of education. Early ethics education in academic settings is critical to prepare future scientists to integrate ethical considerations into their work – including future research in neuroscience. Professional development for experienced investigators is equally important and can serve multiple ends, contributing not only to their individual knowledge, but to the knowledge of the students and young scientists that they mentor as well.” (11, p. 44)

Some of this work has already begun – the Neuroethics Working Group has hosted workshops on invasive and noninvasive neurotechnology involving human research participants as well as the aforementioned NIH Neuroethics Working Group Guiding Principles. The BRAIN Initiative has also established an internal NIH neuroethics program team and has offered neuroethics grants, including fellowships for postdocs, but has not yet developed training grants. Additional training grants would provide an opportunity for more formalized neuroethics training as well as for setting up exportable models of training for graduate students and postdocs. Such training grants might also set up a mentoring cascade in which faculty train/mentor postdocs who then train/mentor graduate students who then train/mentor undergraduate students.

Professional/Institutional support
Since the field of neuroethics is relatively young (about 15 years old, but with a growing community of experts), new approaches are needed to capture talent while also nurturing existing neuroethics scholars – and cultivating cross-fertilization with neuroscience experimentalists. The cultural shifts required to achieve these goals require both top-down and bottom-up methods. Thus, explicit funding opportunities for neuroethics research and for interdisciplinary scholarship are both essential. Research partners are most likely to contribute fully when each is considered an equal participant in the design and conduct of the research – not an “add-on” that is expected to volunteer expertise. To date, the BRAIN Initiative has awarded two rounds of neuroethics research project grants (R01s), and it has included/expanded neuroethics language in predoctoral and postdoctoral training programs. This research has just begun and as awareness of these unique opportunities increases, the program. BRAIN-Initiative support for neuroethics grants should be further continued and expanded in order to anticipate future issues and challenges in BRAIN research as the science progresses.

Committing resources

As a public, taxpayer-funded investment, the BRAIN Initiative aims to promote innovative fundamental science and has a responsibility to assure that the research will be done with integrity and adheres to the highest ethical standards. The BRAIN 2025 report mentions neuroethics as a means to “maximize value” of the neuroscience research investment. In the first few years, leadership of the BRAIN Initiative has increasingly emphasized neuroethics as central – it is our conclusion that this emphasis should not only remain but could also grow over the course of the second half of the BRAIN Initiative. A renewed commitment from the BRAIN Initiative to neuroethical principles amid this ongoing work requires sufficient, dedicated resources to ensure scientific and ethical rigor.

For comparison, the Human Brain Project, (HBP) another similar, large research effort dedicates about 4 percent of its budget to ethics projects – similar to the proportion allocated for ethics in a wide array of biomedical investigations (as in the commitment of up to 5 percent for ELSI research in the Human Genome Project). Applying this focus, from its inception, the HBP continues to conduct a sophisticated and interdisciplinary ethical, societal, and philosophical exploration of how neuroscience could and would inform the question, “What makes us human?” This research endeavor has created a number of sophisticated mechanisms for neuroethics integration and partnerships between ethicists and scientists. For example, the HBP has an ethics advisory board, and each member of this board is partnered with a designated scientist from each project, an “ethics rapporteur.” The board and rapporteurs meet regularly to discuss updates about ethical concerns. The HBP’s Ethics and Society subproject features prominently at the annual HBP meeting and as part of the organization’s progress review.

While the BRAIN Initiative is well underway and now into its second phase, there is still significant opportunity to create additional formal mechanisms to enhance neuroethical inquiry in neuroscience research that will last beyond the formal structure of the BRAIN Initiative. These may include but are not limited to:

- In parallel with the importance of ethics in the Human Genome Project, BRAIN funding for neuroethics should be increased from its current 1.8% to a 5% of the annual budget
Using career-development awards to help support neuroethics researchers
Employing institutional awards to stimulate hiring people with neuroethics expertise
Funding grants for brain science akin to the NIH Centers of Excellence in Genomic Science
Associating center awards supporting neuroethics research with researchers; look at NIH portfolio for opportunities (the Neuroethics Working Group already does this)
Including neuroethics attention/training in relevant training grants
Require a neuroethics section on each BRAIN application, in which the applicant describes the neuroethical issues raised by the proposed research.
Facilitating the matching of a bioethics mentor on certain BRAIN projects to recognize and integrate neuroethics issues would be recognized and brought to the fore and enhance the project

Next-generation focus

Neuroscientists need knowledge beyond what they receive during scientific training to be able to recognize neuroethical issues as well as to conceive neuroethical inquiry in consultation with focused neuroethicists. Focusing on trainees and the next generation of leaders in neuroscience is already happening in the biomedical arena. For example, in collaboration with the International Brain Initiative, professional societies such as the International Brain Research Organization and the Institute of Electrical and Electronics Engineers have partnered with neuroethicists on neuroethics-focused educational modules for in-person and online learning. The International Brain Initiative’s Neuroethics Workgroup is currently designing a neuroethics short course to be shared and offered across the seven existing and emerging large-scale brain projects. Given the vitality of neuroethics training and awareness to the BRAIN Initiative, NIH and other BRAIN-Initiative partners should consider adding additional neuroethics training opportunities within existing responsible conduct of research (RCR) training requirements for neuroscientists. Alternatively, or in addition, the BRAIN Initiative could offer neuroethics-training opportunities associated with funded research at both the trainee- and established-investigator levels, some of which is already being done via administrative supplements. The Neuroethics Working Group has also held several public workshops at which experts considered issues related to BRAIN Initiative-funded research on invasive and noninvasive neural devices (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6010307/).

Beyond the formal structure of the BRAIN Initiative, scientists might work with their local institutions to develop in-house programming featuring integrated neuroethics discussions. Such forums could generate exportable models for informal and formal neuroethics education. Several dedicated neuroethics centers and programs throughout the world, including many in the United States, have modeled undergraduate- and graduate-student neuroethics training and also host neuroethics short courses and regular programming. Some of these institutions offer neuroethics in the context of interdisciplinary training, while others have dedicated neuroscience-training programs. The BRAIN Initiative could offer incentives to academic institutions to offer neuroethics training for neuroscientists, or to join neuroethics training programs between neuroscience and humanities departments, bolstering neuroethics as a part of the neuroscientific enterprise. One opportunity for research might be to survey these institutions for successful strategies for developing neuroethics training programs for neuroscientists. For example, in one case, an undergraduate neuroscience-and-society course offered
to neuroscience majors not only increased knowledge of neuroethics, but also improved overall moral judgment and reasoning skills (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4380300/). Additional research and evaluation would be valuable to inform future programming and to identify the full benefits of neuroethical education for neuroscientists. Given the international make-up and reach of neuroscience research, research within (and sponsored by) the United States benefits from multinational and multicultural participation and leadership. This point is especially important given the need for cross-cultural neuroethics educational models that acknowledge the varied cultural aspects of both ethics and science.

**Global stage for neuroethics**

*What do we know?*

In an era of global science/neuroscience, publishing, and communication – in which both knowledge and the fruits of science transcend geographic boundaries – it has become increasingly apparent that addressing a variety of value frameworks and perspectives is essential for fulfilling the goals of the BRAIN Initiative. Ethical values, assumptions about the role of science, and the types of science that should be pursued actually dictate what science gets pursued. This pattern has become clear in the case of the not-so-gradual move of most NHP research outside of the United States and Europe. Differing values about the conduct of research – along with which and how much data can be collected – have a profound impact on collaboration and data sharing. Questions such as NeQN2 and Guiding Principle 3 (see Chapter 1. Neuroethics: Past, Present, and Future) encourage researchers to carefully consider standards of data collection, as well as to consider potential violations of neuroprivacy.

*What could we learn? Neuroethics research opportunities*

NeQN3 encourages scientists to consider ethical issues that arise from innovative models of neural circuitry. One example is that posed by brain organoids/assembloids that are genetically engineered to model human brain development, cortical regions, and diseases. While closer approximations to human brains afford richer opportunities to gain deeper insights into the human brain and behavior, these models will also raise concern about the appropriateness of their use given their similarity or similar capacity to human brains. A similar debate has arisen in the context of the use of CRISPR-modified NHPs to study autism (https://www.ncbi.nlm.nih.gov/pubmed/?term=30329048).

It is also important to consider the inevitable use of lab-generated technologies for purposes beyond their original intent. (see Chapter 5 of this report). This possibility is recognized in Guiding Principles 4 and 5 (see p. see Chapter 1. Neuroethics: Past, Present, and Future), but needs additional attention. In contrast, the Human Brain Project has a sophisticated network of neuro ethicists who have collaborated with scientists to create a variety of opinion-pieces (https://www.sciencedirect.com/science/article/pii/S0896627319300066?via%3Dihub) on key topic areas. These international groups (sometimes including individuals from up to 20 countries) accomplish the difficult task of harmonizing and reconciling differing views.

One recent Human Brain Project Opinion covered the topic of dual-use research, referring to uses and applications of research beyond the initially conceived hypothesis (mentioned within NeQN5 and in Guiding Principle 5). These Opinions are drafted and published by the Human Brain Project with input...
from an interdisciplinary group of ethicists, philosophers, and social scientists, including from the project itself. These well-researched reports are generally informed by both science as well as public-engagement research on specific topics.

NeQ2 asks researchers to explore ethical standards of biological material and data collection as well as how they relate to those of global collaborators. As the culture around data collection is moving toward one of sharing and openness, researchers around the globe will need to be aware of also-shifting tides of acceptability and regulation of non-human animal research, particularly as these models attempt to become closer approximations of human disease and suffering (https://www.ncbi.nlm.nih.gov/pubmed/?term=30329048).

Public engagement: Meaningful and bidirectional

The modern consensus on how to approach and achieve public engagement for scientific pursuits is quite different from past strategies that focused on increasing public knowledge of science. The latter, mostly unidirectional methods mirror the information-deficit model of science communication – a model that has fallen from favor in both the science-communication and educational communities. Instead is the recognition that individuals within the public arena make conscious choices about what they want to know and learn, as well as how those efforts align with personal and societal values.

What do we know?

Both scientists in training and non-scientists alike take great interest in neuroscience, based upon the anticipation that advances and discoveries in brain research will affect how we understand ourselves as well as how we engage with the world. Neuroethics is thus a common entry point to neuroscience for everyone. Meaningful public engagement is critical to the success of neuroscience, as articulated by two of the BRAIN Initiative’s Neuroethics Guiding Principles:

- Principle 6: Identify and address specific concerns of the public about the brain
- Principle 7: Encourage public education and dialogue

The BRAIN Initiative also communicated the importance of public involvement and engagement, “Stakeholders should be engaged through a variety of additional mechanisms, including academic research in bioethics, training programs for a broad array of practitioners and students in the medical professions, conferences targeted to audiences with different levels of scientific expertise, and media outreach.” The return on investment from publicly funded research rests on the strength of the public's trust in individual scientists and with the scientific enterprise. Like many new technologies and scientific advances, neuroscience advances are frequently subject to hyperbole. Importantly, we cannot only blame the media for such hype. Scientists must appreciate their own responsibility to communicate their work to general audiences clearly and effectively – while retaining its genuine interest and excitement.

What could we learn? Neuroethics research opportunities

Communicating science with non-scientists via a deficit model that assumes the public is wholly ignorant of science is not only dismissive but also unlikely to be successful (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5753839/). Skilled communication and effective
engagement will likely require resources to connect scientists with experts in public engagement. While it is not uncommon for public engagement to be explicitly required as a component of conducting research projects, rarely are sufficient resources devoted to rigorous interdisciplinary collaborative work (https://www.oecd-ilibrary.org/science-and-technology/neurotechnology-and-society_f31e10ab-en) in this area. Scientists of tomorrow (and today) must be prepared to address the reality that science is being communicated, formally and informally, through a relentless 24-hour, 7-day media cycle. In summary, scientists should learn to be adept at public scholarship and engagement.

Some important considerations for the BRAIN Initiative related to engagement opportunities, particularly in neuroethics, include principles and lessons learned from the societal experience with human-genome editing. As noted in the 2017 National Academies of Science, Engineering, and Medicine’s Human Genome Editing: Science, Ethics, and Governance:

“We need to engage the public in a more open and honest bidirectional dialogue about science and technology and their products, including not only their benefits but also their limits, perils, and pitfalls. We need to respect the public's perspective and concerns (https://www.ncbi.nlm.nih.gov/pubmed/12586907) even when we do not fully share them, and we need to develop a partnership that can respond to them.

The authors of this report noted that high-quality engagement is marked by systematic exploration of the full range of risks and benefits of technology that go beyond simply those that are technical and medical, but that include perspectives and knowledge from all interested and affected parties (https://www.nap.edu/read/5138/chapter/1). Other important considerations include assessing quality outcomes from engagement work that explores policy and regulatory issues – those that consider both facts and values, as well as how anticipated societal effects will affect the things people value (https://www.nap.edu/read/12434/chapter/1). Legitimate engagement practices are those participants view as transparent, fair, and competent (https://link.springer.com/chapter/10.1007/978-94-011-0131-8_13), and which truly interrogate the values and interests of the people who will be using these technologies or otherwise affected by them. On a practical note, engagement work cannot be done successfully without administrative efficiency, which requires dedicated resources. To this end, the BRAIN Initiative could consider supporting neuroethics research that assesses public opinion as well as develops best practices for public engagement around neuroethics issues.

Many challenges are inherent in attempting public engagement in the modern world. We live in a global society in which information access and spread is rapid and distributed – sometimes without proper context. Thus, it is critical to consider engagement activities internationally, requiring a broad definition of stakeholders that extend beyond English-speaking countries. This is especially important given that individuals across the world may use neurotechnologies. The Internet, social media, and other creative electronic and in-person formats are powerful tools for public engagement, but they carry significant risk for distributing unvetted information and/or unsubstantiated claims, in a manner that is difficult to control.

A particularly vexing challenge is translating outcomes of science and related engagement activities into changes in policy and practice. Success requires controlling two key levers: i) support from those empowered to make decisions to incorporate public views and values; and ii) transparent, justifiable,
and monitored pathways for those actions. Bias and ulterior motives can also be a concern: What about when decision makers use engagement activities toward achieving predetermined outcomes? In summary, effective public engagement is highly collaborative and requires input from individuals and groups responsible for funding, doing, and measuring such activities.

There are solid examples of what works, such as public-expert interactions that can produce meaningful communication about neuroscience as provided by the National Information STEM Education (NISE) Network, which balance understanding and engagement for various topics, assimilating and integrating the different ways public and scientific individuals and audiences interact. The NISE Network published a 2018 conference report, “Public Engagement with Neuroscience and Society (http://nisenet.org/sites/default/files/nise_net_kavli_conference_report_1.pdf),” which notes that sustained public engagement will benefit from leveraging existing strengths of the current neuroscience outreach ecosystem. Components include BRAIN Initiative-funded neuroscience research, a comprehensive educational content framework (e.g., BrainFacts from the Society of Neuroscience), museums with broad reach that use evidence-based engagement approaches, and employing as ambassadors volunteer experts (e.g., Brain Awareness Week and activities therein).

Next steps: implementable goals

Integrating neuroethics and neuroscience is happening, but to fully reap the benefits of the BRAIN Initiative, closer alignment is needed to ensure scientific and ethical rigor – and also to share both the sense of amazement and practical outcomes from this groundbreaking large-scale, cross-sector project. Key concepts toward achieving this goal include fostering side-by-side professional interactions between neuroethicists and neuroscientists, extending to neuroscientists knowledge and appreciation of neuroethical principles embedded in basic neuroscientific inquiry, offering formal neuroethics training at various career levels, continuing to support neuroethics research, and truly embracing public engagement as an opportunity to fortify the research investment. Specific possibilities include:

- Establishing (continue to offer, via supplemental funding and neuroethics R01s) and expanding formal mechanisms and incentives to embed (neuro-)ethicists within neuroscience research projects
- Supporting trainees and the next generation of leaders in neuroscience and neuroethics
- Establishing formal opportunities for established scientists and trainees to learn about neuroethics – and for neuroethicists to learn about neuroscience
- Using published principles and guidelines such as the NeQNs and Guiding Principles to provide both scientists and neuroethicists a springboard to discuss the design, conduct, and translation of neuroscience research
- Establishing a neuroethics network resource, consisting of people to consider issues on an ongoing basis for a range of stakeholders (neuroscience researchers and trainees, IRBs, health care providers, non-scientific public)
- Developing NIH BRAIN mechanisms for institutional support and incentives to conduct collaborative, interdisciplinary neuroscience research, perhaps through funding centers and/or joint applications
• Create training grants and other funding strategies to explore more formalized neuroethics training, which may also yield exportable models of training for graduate students and postdocs
• Partnering with public-engagement experts, including those using innovative methodologies
• Investigating the relevant neuroethical concerns of BRAIN investigators and of the public
• Developing and evaluating neuroethics educational programs and assessments
• Identifying successful strategies and models for effective neuroethics engagement

NEUROETHICS MOONSHOT - Revolutionizing BRAIN: The Theory of the Mind

APPENDIX 1: ROSTER

Priority Areas
Funding Opportunities

Cell Type (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=1)

Circuit Diagrams (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=18)

Monitor Neural Activity (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=2)

Interventional Tools (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=3)

Theory & Data Analysis Tools (/funding/funding-opportunities?combine=&field_priority_area_tid%5B%5D=4)
Human Neuroscience (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=5)

Integrated Approaches (/funding/funding-opportunities?
combine=&field_priority_area_tid%5B%5D=19)

BRAIN Update (https://www.brainupdate.nih.gov/)

BRAIN 2025 Report (/strategic-planning/brain-2025-report)

BRAIN Initiative Alliance (https://www.braininitiative.org/)

BRAIN Scientific Priority Areas (/strategic-planning/brain-priority-areas)
NIH Institutes & Centers


BRAIN Initiative Partners

Alliance Members and Affiliates (http://www.braininitiative.org/alliance/)
All Participants (http://www.braininitiative.org/participants/)

More Information

Email Us (/about/contact-us)
Archives (/archives)
BRAIN Update Blog (https://brainupdate.nih.gov/)

Related Websites

National Institutes of Health (https://www.nih.gov)
USA.gov (https://www.usa.gov)

Other Links

Adobe Acrobat (https://get.adobe.com/reader/)

Email Us (/about/contact-us)  Archives (/archives)  BRAIN Update Blog (https://brainupdate.nih.gov/)

SCANNING FOR JUSTICE: USING NEUROSCIENCE TO CREATE A MORE INCLUSIVE LEGAL SYSTEM

Hilary Rosenthal*

ABSTRACT

Although they may seem to be worlds apart, on further inspection, neuroscience and the law are not so discordant. Neurolaw is an emerging interdisciplinary field that undertakes to examine how an increased understanding of the human nervous system can lead to a more precise explanation for human behavior, which in turn could inform the law, legislation, and policy. While increased dependence on neuroscience in the courtroom raises evidentiary and normative concerns, its use can also have significant implications for civil and human rights by opening doors for plaintiffs to bring claims that historically have been difficult to prove. One such example is the way neuroscience can obviate the outmoded physical-mental divide in tort law. Courts in the United States have been skeptical of awarding damages for “invisible” injuries, such as PTSD, concussions, neurodegenerative diseases, and emotional pain and suffering, all of which can alter brain structure and function, but often do not manifest physically until it is too late for a person suffering those harms to recover damages in a courtroom. However, as neuroscience technology improves, it can help detect these previously hidden or latent injuries, especially for those in marginalized communities, and begin to uproot entrenched policies that perpetuate health inequality. This Note argues that neuroscience, while not without its shortcomings, has become an

* J.D. Candidate 2019, Columbia Law School; B.A. 2012, Brown University. The author would like to thank Professor Kristen Underhill for her exceptional guidance on this piece, the staff of the Columbia Human Rights Law Review, especially Ruth O’Herron, for their invaluable assistance, and the author’s peers at Columbia Law School for their unwavering support.
increasingly important tool to create a fairer, more just, and more rehabilitative justice system.
# Table of Contents

Table of Contents ................................................................. 292  

Introduction ............................................................................... 293  

I. Neurolaw: What It Is and How It Is Used .................... 296  
   A. Types of Neuroscience Technology .......................... 297  
   B. The Intersection of Law and Neuroscience .......... 300  
   C. Neuroscience as a Burgeoning Tool in Tort and Human Rights Claims ............................. 304  

II. Neurolaw’s Potential to Advance Human Rights in Tort Cases ................................................................. 306  
   A. Tort Law and the Decline of Substance Dualism ...... 306  
   B. Neuroscience and the Identification of Pain .......... 311  
   C. Neuroscience and PTSD ......................................... 314  
   D. Neuroscience’s Potential to Identify Mild Traumatic Brain Injury ............................................... 317  
   E. Neuroscience’s Potential to Identify Neurotoxins in the Brain and Nervous System .................. 321  

III. Challenges to and Weaknesses of Neuroscience’s Use in the Law ................................................................. 326  
   A. Evidentiary Concerns ............................................. 327  
   B. Normative, Ethical, and Policy Challenges to Neuroscience in the Courtroom ..................... 331  

Conclusion ................................................................................ 337
INTRODUCTION

“The age of scanning has dawned in our courtrooms. This is not a technological genie we are going to be able to put back in the bottle.”

Imagine you are a young child on your way home from your first day of the second grade. You are riding on the school bus, and are excited, if a little nervous, because it is your first time riding on your own. It has been a long day, and you accidentally fall asleep and miss your stop. You awake later to find yourself in an empty bus, parked in a lot nowhere near your home. For many of us, this might incite feelings of fear, anxiety, and distrust. You might be afraid of going to school the next day, perform poorly if you do, or even become unwell when faced with the prospect of getting on a bus again.

If you brought an action in negligence against the bus driver for an incident like this, it might be difficult to prove in court just how much the experience distressed you. You may not have any outward physical symptoms, but anyone who has had an experience like this can tell you that the lasting psychological effects do not easily fade from the mind. What if those processes going on in your brain causing you to feel fear, anxiety, or distrust could be mapped and shown to prove your distress?

For a young boy named Daniel, this kind of evidence could have been helpful. A psychologist diagnosed Daniel with post-traumatic stress disorder (hereinafter PTSD) after a traumatic experience of being abandoned on a school bus. While the court acknowledged his distress as a real, debilitating injury, it did not find that it qualified as a “physical injury” as required under Kansas tort law and, therefore, Daniel could not recover damages. But what if there was a way to show that the psychological effects from which Daniel was suffering were not based in some existential ether, but had physiological roots in the brain?

The field of cognitive neuroscience has the potential to do just that. Recent advancements in this field should change the way we

3. Id. at 619.
think about the physical-mental divide in tort law. As a judge wrote in dissent in Daniel’s case,

[R]esearch does not support a categorical distinction between emotional and physical harm . . . "Whatever the best minds of the day might have thought about the difference in physical and emotional harm when tort law came of age, the best minds of today do not support such a stark mind-body dichotomy." 4

Severe emotional distress is just one type of injury that people all across the world suffer but may not visibly display, even though it can significantly affect one's daily life. Other injuries, such as concussions and neurodegenerative diseases, can also alter brain structure and function, but often do not manifest physically until it is too late for a person suffering those harms to recover damages in a courtroom. 5 Courts in the United States have been skeptical of awarding damages for these injuries for evidentiary and normative purposes, as in Daniel’s case. Historically, there has been a lack of adequately objective evidence for claims that we cannot see with the naked eye. In a legal system where resources are limited, courts have required some sort of dividing line to determine harms worthy of compensation. But as technology that analyzes the brain and its functions improves, we can start identifying these previously hidden or latent injuries and rectify gaps for redress in tort and other bodies of law.

This Note focuses primarily on tort law as an essential mechanism for enforcing civil rights in a common law system and a means by which citizens can maintain their human rights on a transnational scale. These rights include, among others, the rights to health, 6 equality before the law, 7 and dignified treatment, 8 as well as a

4. Id. at 621 (Green, J., dissenting) (internal citations omitted) (quoting Daniel W. Shuman, How We Should Address Mental and Emotional Harm, 90 Judicature 248, 248 (2007)).
5. By the time the damages are identified, a victim might be beyond repair. See, e.g., Emily Kelly, I'm the Wife of a Former N.F.L. Player. Football Destroyed His Mind, N.Y. TIMES (Feb. 2, 2018), https://www.nytimes.com/2018/02/02/opinion/sunday/nfl-cte-brain-damage.html (on file with the Columbia Human Rights Law Review) (chronicling just one of many accounts of former professional athletes suffering permanent brain injuries).
7. Id. at art. 7.
8. Id. at art. 1.
right to due process and to a fair trial in a domestic setting. neuroscience has become increasingly important in ensuring these rights are protected, and may prove useful as a tool to create a fairer, more just, and more rehabilitative legal framework.

This Note argues that, despite some very serious evidentiary and normative concerns, as neuroscience technology becomes increasingly accurate, less expensive, and more precise in illustrating the ways in which people suffer harm, traditional dichotomies of injury compensation, such as the physical-mental divide, are no longer valid. Instead, the law should compensate based on severity of harm rather than type of injury. Redefining this line in tort law will uphold and advance individual autonomy and normative values inherent in our tort system, resulting in more accurate and objective compensation that utilizes modern technology to help people who would otherwise go without just compensation.

Part I of this Note provides a general overview of the intersection between neuroscience and law and the different technologies involved in examining injury, pain, and emotion in the brain. It then details some of the ways that neuroscience technology is already used or discussed in courtrooms and in legislation. Part II describes how neuroscience technology can be used to advance human and civil rights, particularly in tort law, by helping plaintiffs receive compensation for invisible injuries that previously have been difficult to prove, focusing on case studies of PTSD, mild traumatic brain injuries, and neurotoxicity. Part III of this Note explains a few of the ways that neuroscience technology and its use in the courtroom could backfire or hurt some litigants. It also describes some complications of this technology and explains how it can nonetheless be useful if certain precautions are taken. Part III also suggests a few procedural options for regulating or evaluating neuroscience evidence in civil courtrooms.

12. Adam J. Kolber, Will There Be a Neurolaw Revolution?, 89 IND. L.J. 807, 822 (2014) (arguing that neuroscience technology has become increasingly reliable. For example, functional magnetic resonance imaging scans can, at least in controlled experimental contexts, predict with 80% accuracy whether or not a particular subject is in pain).
advocating in particular for a working group that could research, draft, and oversee policy proposals. This Note ultimately argues that, as brain scanning technology becomes increasingly precise and research surrounding it increasingly refined, lessons learned from neuroscience will inevitably influence the law and that, overall, inclusion of neuroscientific evidence in the courtroom is advantageous and desirable for tort litigants whose claims until recently have been unverifiable.

I. NEUROLAW: WHAT IT IS AND HOW IT IS USED

At first glance, neuroscience and the law may seem to be odd bedfellows. However, an increasing amount of scholarship and attention is being addressed to the ways that neuroscience—the scientific study of the structure and function of the nervous system and the brain—may impact law, legislation, and policy.\(^1\)\(^4\) Neuroscience research is rapidly developing and illuminating our understanding of human behavior, motivation, intention, and cognition.\(^1\)\(^5\) Understanding how our brains function from a physiological viewpoint should affect how we think about and normatively construct the law. Because neuroscience is the study of the cognitive processes that underlie human behavior, it ought to have significant implications for legal systems, which are ultimately concerned with the regulation of human behavior. Indeed, it already has.\(^1\)\(^6\)

This Part discusses the intersection of neuroscience and law. Section A describes the neuroscience technologies involved in

14. See, e.g., Kolber, supra note 12, at 808 ("[T]here will indeed be a neurolaw revolution. It may arise . . . from a wave of new brain technologies that will change society and the law in a wide variety of ways.").
16. There are increasing signs of neuroscience technologies becoming less of a science fiction plot and more of a reality. For example, in the United States, at least two companies, No Lie MRI and Cephos Corp., have offered magnetic resonance imaging-based lie-detection services. Richard Birke, Neuroscience and Settlement: An Examination of Scientific Innovations and Practical Applications, 25 OHIO ST. J. ON DISP. RESOL. 477, 482–83 (2010); Eli Aharoni et al., Neuroprediction of Future Rearrest, 110 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES 6223 (2013).
examining injury, pain, and emotion in the brain. Section B provides background and a general overview of the use of neuroscience in legislation, policy, and the courtroom. Section C addresses some of the ways that these technologies can be used in tort and human rights law specifically, focusing on how they can substantiate the invisible injuries affecting people across the world.\(^ {17} \)

A. Types of Neuroscience Technology

Among the various neuroscience technologies used to examine injury, pain, and emotion in the brain, courts and scholarship to date have focused mostly on functional magnetic resonance imaging (hereinafter fMRI) and positron emission tomography (hereinafter PET) scans and how they reflect the physical processes that take place in the brain.\(^ {18} \)

An fMRI measures blood oxygenation levels in the brain and enables scientists to detect which brain regions are receiving more blood flow.\(^ {19} \) When there are changes in brain activity, such as when a patient feels a painful stimulus, blood flow throughout the brain changes as hemoglobin in the blood carries oxygen to the areas of the brain that are working harder.\(^ {20} \) When the hemoglobin releases oxygen to those areas, it becomes paramagnetic,\(^ {21} \) which triggers a magnetic

\(^ {17} \) Notably, the Office of the United Nations High Commissioner for Human Rights recognizes the “human right of everyone to the enjoyment of the highest attainable standard of physical and mental health” (emphasis added). See Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health, U.N. Doc. A/HRC/35/21 (March 28, 2017).

\(^ {18} \) These technologies in particular can illustrate the physical processes occurring in our brains in real time. Our brains hold about 100 billion neurons, each one making 1,000 or more connections—"synapses"—to other neurons in which they are constantly giving off and picking up chemicals called neurotransmitters, which communicate information throughout the brain and body. Henry T. Greely, Neuroscience, Mindreading, and the Courts: The Example of Pain, 18 J. OF HEALTH CARE L. & POL’Y 171, 171 (2015).


\(^ {20} \) Id. at 5.

\(^ {21} \) Deoxyhemoglobin, present in deoxygenated blood, is paramagnetic; that is, its presence causes a decrease in a magnetic resonance signal. O. Carter Snead, Neuroimaging and the “Complexity” of Capital Punishment, 82 N.Y.U. L. REV. 1265, 1285 (2007).
field located inside a magnetic resonance imaging (MRI) scanner. Neuroscientists can track these signals, referred to as the BOLD (blood-oxygen-level dependent) signals, and see how they flow to different areas of the brain over time.

PET scans also measure blood flow in the brain. PET researchers inject a radioactive tracer into the bloodstream and, by tracking its path, can identify neural brain activity in particular areas of the brain. PET and fMRI scans therefore identify the portions of the brain that are activated when a person is experiencing or thinking something, based on the increased quantity of freshly oxygenated blood the regions draw. During a scan, a researcher can measure and correlate the brain areas receiving more blood flow at different time intervals as a participant is asked questions or given a stimulus, such as pain.

Other types of neuroscience technology used to measure abnormalities or disorders in the brain include single-photon emission computed tomography (SPECT), electroencephalography (EEG), quantitative electroencephalography (qEEG), and magnetoencephalography (MEG) scans. These various types of scans can be used, respectively, to distinguish between different types of seizures, to pinpoint defects in auditory and somatosensory areas, to diagnose sleep disorders, and to examine head injuries, tumors, infections, and neurodegenerative diseases. These neuroimaging methods are largely non-invasive, safe, increasingly accessible, and less expensive than other types of scans, such as fMRIs.

22. Id. at 7.
25. Id. In numerous countries, including the United States, the prevalence of fMRI machines per population has been steadily increasing. Number of magnetic resonance imaging (MRI) units and computed tomography (CT) scanners: Selected countries, selected years 1990–2009, CTRS. FOR DISEASE CONTROL & PREVENTION (2011), https://www.cdc.gov/nchs/data/hus/2011/123.pdf [https://perma.cc/Q9VH-7494].
26. Greely, supra note 18, at 190 ("EEG . . . has many advantages over fMRI—it is cheap, portable, and easy to operate.").
27. Although fMRI scans may be prohibitively expensive for many plaintiffs. See Stephen J. Morse, Neuroimaging Evidence in Law: A Plea for
However, it is important to keep in mind that functional brain imaging is not necessarily the same thing as mind reading. Researchers warn that "[w]hile fMRI can accurately measure changes in blood flow and oxygen levels, interpreting those changes as reliable indicators of particular types of thought, or as reliable indicators of what a region of the brain is actually doing, requires a series of inferential steps that are not entirely straightforward."\textsuperscript{28} Perhaps even more critical is that this technology cannot prove causation. Although differences in brain scans after an emotionally disturbing event may have a corollary relation, this does not mean that the emotionally disturbing event was the cause of those differences, and other events may have contributed to the injury since. Similarly, when trying to image how a plaintiff's pain or injury has changed over time, researchers may need a baseline scan of the plaintiff's pain level before a defendant's wrongful action, which is often absent.\textsuperscript{29}

Despite these weaknesses, neuroscientific evidence from fMRIs, PET scans, and other tests is increasingly used in U.S. courtrooms under certain evidence standards,\textsuperscript{30} and the rapid growth of neuroscience technology will likely only improve its accuracy and reduce its price.


29. Floyd Bloom et al., \textit{Does Neuroscience Give Us New Insights into Drug Addiction?}, in \textit{A JUDGE'S GUIDE TO NEUROSCIENCE: A CONCISE INTRODUCTION} 34 (2010), \url{https://www.sagecenter.ucsb.edu/sites/staging.sagecenter.ucsb.edu/files/file-and-multimedia/A_Judges_Guide_to_Neuroscience%5Bsample%5D.pdf} ([https://perma.cc/L2L7-SQT2] ("To determine the pattern of activity, the fMRI BOLD signal during pain has to be compared with a baseline condition when there is no pain").

30. See \textit{FED. R. EVID.} 401, 702; Daubert v. Merrill Dow Pharm., Inc., 509 U.S. 579, 589 (1993) (holding that trial judges must determine whether expert testimony is both "relevant" and "reliable"); Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923) (holding that expert testimony must be based on knowledge that has "gained general acceptance in the particular field"). The \textit{Daubert} standard is the law in federal court and over half of the states, while the \textit{Frye} standard is preferred in some jurisdictions including California, Illinois, Maryland, New Jersey, Pennsylvania, and Washington. Amanda C. Pustilnik, \textit{Imaging Brains, Changing Minds: How Neuroimaging Can Transform the Law's Approach to Pain}, 66 \textit{ALA. L. REV.} 1099, 1148 (2014).
B. The Intersection of Law and Neuroscience

The intersection of law and neuroscience, often dubbed "neurolaw," has featured prominently in a number of cases and policy debates in the U.S. Scholars, advocates, and judges have invoked neurolaw in a variety of legal fields, including criminal responsibility and sentencing, lie detection, adolescent brain development and

31. "Neurolaw" is the application of neuroscience findings to legal topics such as criminal, tort, and administrative litigation and justice, agency, intent, and policy. Neil Aggarwal & Elizabeth Ford, The Neuroethics and Neurolaw of Brain Injury, 31 BEHAV. SCI. L. 789, 790 (2013).

32. See generally Francis X. Shen, The Overlooked History of Neurolaw, 85 FORDHAM L. REV. 667, 1043-49 (2016) (discussing how the intersection of neuroscience and the law is centuries old). The number of cases in the U.S. involving neuroscientific evidence doubled from 2006 to 2009 and there are a growing number of criminal cases involving neuroscientific evidence. Jones & Shen, supra note 28, at 353.

33. In State v. Nelson, F05-846 (Fla. 11th Cir. Ct. 2010), qEEG evidence was admitted into evidence in a U.S. court for the first time, and contributed in part to the jury voting to sentence the defendant to life in prison instead of the death penalty. One juror commented that "the technology really swayed me . . . after seeing the brain scans, I was convinced this guy had some sort of brain problem." David Ovalle, Novel defense helps spare perpetrator of grisly murder, MIAMI HERALD (Dec. 2, 2010), http://www.floridacapitalcases.state.fl.us/Documents/Enewsletter/2010_Articles_December/Novel%20defense%20helps%20spare%20perpetrator%20of%20grisly%20murder.pdf [https://perma.cc/KJ3D-RMG7].

34. There have been various instances in which fMRI and EEG-based lie detection evidence was proffered in U.S. courts. United States v. Semrau presented the first evidentiary hearing in federal court on the admissibility of fMRI lie-detection evidence. 693 F.3d 510, 521 (6th Cir. 2012). Although Magistrate Judge Pham ultimately did not admit the evidence under Federal Rule of Evidence 702, he wrote that "in the future, should fMRI-based lie detection undergo further testing . . . this methodology may be found to be admissible." Amended Report and Recommendation, United States v. Semrau, 2010 WL 6845092, at *12 n.18 (W.D. Tenn. June 1, 2010).
juvenile justice, morality, free will, risk and information processing in addicts, brain death and injury, judge and jury bias, and tort law, among others. The use of neuroscience data as evidence in U.S. courtrooms has risen sharply in the past decade. Outside of the courtroom, neuroscience is also widely cited in various types of legislative bills, especially relating to brain injury.


36. Neuroscience has shed light on some of the ways people think about moral conundrums, such as the classic trolley hypothetical. Experimenters found that the brain region associated with deliberate problem solving and self-control, the dorsolateral prefrontal cortex, was especially active in an fMRI scan when subjects contemplated the utilitarian option of pulling the lever and saving the greatest number of lives. By contrast, the ventral medial prefrontal cortex, an emotional center of the brain, was active when subjects imagined harming the one individual even if it would have saved others. Jeffrey Rosen, *The Brain on the Stand*, N.Y. TIMES MAG. (Mar. 11, 2007), http://www.nytimes.com/2007/03/11/magazine/11Neurolaw.t.html (on file with the Columbia Human Rights Law Review).

37. See Steven K. Erickson, *Blaming the Brain*, 11 MINN. J.L. SCI. & TECH. 27, 28 (2010) (evaluating how cognitive neuroscience research may disrupt longstanding norms of personal responsibility).


40. Lisa G. Aspinwall et al., *The Double-Edged Sword: Does Biomechanism Increase or Decrease Judges’ Sentencing of Psychopaths?*, 337 SCIENCE 846, 846 (2012). One study found that certain extraneous variables, such as the time of day, affect a judge’s decision to grant parole. Areas of the brain that are used for more complex reasoning, such as the dorsolateral prefrontal cortex, show less activity when other areas concerned with more basic bodily needs like hunger, such as the temporal lobe, are more active. Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT’L ACAD. SCI. 6889, 6892 (2011).


medical insurance,\textsuperscript{44} mental health, education and early childhood interventions, and veterans' affairs.

Neuroscience technology has had varying influence in the criminal courtroom context thus far.\textsuperscript{45} Brain scans have been presented to mitigate a defendant's sentence,\textsuperscript{46} to show that a defendant is incompetent to stand trial, and to prove that a defendant did not have the requisite mens rea at the time of the crime, though the scans are not uniformly admitted or successful. Sometimes the scans merely demonstrate correlation rather than causation—at least in a courtroom's eye. For instance, in \textit{People v. Goldstein}, a defendant who pushed a woman in front of a subway train to her death sought to introduce a PET image of a brain abnormality in an effort to prove an insanity defense of schizophrenia.\textsuperscript{47} While the prosecution conceded that Goldstein suffered from schizophrenia, the court excluded the PET scan from evidence because even though it demonstrated a brain abnormality, it would not actually be probative as to the mens rea element of the crime, “since a diagnosis of schizophrenia does not preclude per se that a defendant is capable of such comprehension.”\textsuperscript{48} Cases such as this one illustrate that, at least in criminal contexts in the United States, neuroscientific data can be useful as an evidentiary tool, though a court may reject the admissibility of the data if it cannot prove a causal basis for a defendant's actions or if it is not used in conjunction with corroborating evidence.

\textsuperscript{44} Stacey A. Tovino, \textit{Will Neuroscience Redefine Mental Injury? Disability Benefit Law, Mental Health Parity Law, and Disability Discrimination Law}, 12 IND. HEALTH L. REV. 695, 697–727 (2015) (exploring the role that brain scanning technologies play in securing health insurance coverage, social security eligibility for mental health conditions, and in officially recognizing gender-specific mental health conditions like premenstrual syndrome and postpartum depression).

\textsuperscript{45} Eggen & Laury, supra note 41, at 238 (“[T]he criminal courtroom has become an early testing ground for the application of the studies to cognitive mental states in the law. The courts have shown interest, tempered by caution, and suspicion of the evidence's reliability.”).

\textsuperscript{46} Introducing evidence to demonstrate brain abnormalities or injuries for the purpose of mitigating sentencing has been one of the more common uses of neuroscience in the courtroom. For example, an Oregon boy convicted of killing and injuring fellow students introduced images showing brain abnormalities and was granted a more lenient sentence due to his mental illness. \textit{State v. Kinkel}, 56 P.3d 463, 467 (Or. Ct. App. 2002).


\textsuperscript{48} \textit{Id.}
Neuroscience evidence has also been used in various cases and legislation concerning juvenile justice. In *Miller v. Alabama*, the United States Supreme Court cited brain science findings regarding impulse control, planning, and risk avoidance in holding that mandatory sentences of life without the possibility of parole are unconstitutional for juvenile offenders. That case, and at least twenty others, have relied on an affidavit written by Ruben Gur, a national PET expert, arguing that adolescents are not as capable of controlling their impulses as adults because the development of neurons in the prefrontal cortex is not complete until the early 20s. Similarly, in *Graham v. Florida* the Supreme Court cited neuroscientific and psychological data on adolescent development when it struck down, under the Eighth and Fourteenth Amendments, sentencing juveniles to life in prison without parole for non-homicide crimes.

Courts have also considered neuroscience evidence in civil matters. In his dissent in *Brown v. Entertainment Merchants Association*, Justice Breyer cited "cutting-edge neuroscience" to

49. For instance, the neuroscience of adolescent development featured prominently in a 2011 California Senate Bill which allowed juveniles sentenced to life without parole to submit a request to have a new sentencing hearing. CAL. PENAL CODE § 1170(d) (West 2011).

50. *Miller v. Alabama*, 567 U.S. 460, 471–73, 472 n.5 (2012) (favorably citing neuroscience evidence presented in amicus briefs, Justice Kagan wrote for the majority, "We reasoned that those [neurological] findings—of transient rashness, proclivity for risk, and inability to assess consequences—both lessened a child's 'moral culpability' and enhanced the prospect that, as the years go by and neurological development occurs, his 'deficiencies will be reformed'").

51. Declaration of Ruben C. Gur, Ph.D. at 15, *Patterson v. Texas*, 536 U.S. 984 (2002), https://www.americanbar.org/content/dam/aba/publishing/criminal_justice_section_newsletter/crimjust_juvjus_Gur_affidavit.authcheckdam.pdf [https://perma.cc/6DA6-88CW] ("The evidence now is strong that the brain does not cease to mature until the early 20s in those relevant parts that govern impulsivity, judgment, planning for the future, foresight of consequences, and other characteristics that make people morally culpable.").

52. *Graham v. Florida*, 560 U.S. 48, 68 (2010) ("[D]evelopments in psychology and brain science continue to show fundamental differences between juvenile and adult minds."). *See also* *Roper v. Simmons*, 543 U.S. 551 (2005), in which the Supreme Court struck down the death penalty for offenders who committed crimes when they were under the age of 18 partially based on an affidavit that argued that because adolescents' prefrontal cortices are not fully developed, they are less able than adults to control their impulses and should not be held fully accountable "for the immaturity of their neural anatomy . . . [because] it is a degree never before understood, scientists can now demonstrate that adolescents are immature . . . in the very fibers of their brains."). Brief for American Medical Association et al. as Amici Curiae Supporting Respondent at *10, Roper v. Simmons*, 543 U.S. 551 (2005) (No. 03-633).
support the argument that violent video games are linked to more aggressive behavior. In other cases, neuroscientific evidence has been successful in proving mental incapacity. For example, in Van Middlesworth v. Century Bank and Trust Co., a defendant introduced brain scans to prove his mental incompetency, resulting in the court ruling that the real estate contract that he signed was void.

C. Neuroscience as a Burgeoning Tool in Tort and Human Rights Claims

Scholars have debated the use of neuroscience evidence in criminal sentencing mitigation at length, but an increased reliance on neuroscience in the courtroom, in legislation, and in regulation can also have significant implications for civil and human rights and can open many new doors for plaintiffs to bring claims, such as in tort law. By pinpointing cognitive responses with objective technology, neuroscience tools and research can provide substantiation for “invisible" tort injuries that historically have been difficult to prove

55. Francis X. Shen, Law and Neuroscience 2.0, 48 ARIZ. ST. U. L. REV. 1043, 1049–50 (2016) (remarking that neurolaw often focuses only on criminal law, and that many of its other intersections with the law are ripe for discussion).
57. Marcello Ienca & Roberto Andorno, Towards New Human Rights in the Age of Neuroscience and Neurotechnology, 13 LIFE SCI., SOCY & POLY (Apr. 26, 2017), at 2, 8 (writing that “international human rights law does not make any explicit reference to neuroscience. In contrast to other biomedical developments ... neurotechnology still largely remains a terra incognita for human rights law”).
58. “Invisible injuries” may be defined as those that affect a person's enjoyment of life but are not visible to the naked eye or do not readily show up on traditional technologies such as X-rays. Some examples include mTBIs, PTSD, chronic pain, fibromyalgia, emotional distress, and neurobiological toxins. See
beyond a plaintiff's testimony, such as PTSD, mild traumatic brain injuries (hereinafter mTBIs), toxic exposure, and emotional pain and suffering. Neuroscience technology's ability to reveal the pain response in an individual's brain could reform how we think about tort law and what compensation may be due to litigants. Incorporating this emerging field into law would be especially beneficial for the rights of people in marginalized communities, such as racial and ethnic minorities,\(^5\) prison inmates, indigent people,\(^6\) abuse victims, and invisible injury victims,\(^6\) whose pain and suffering traditionally is underreported.

PET, fMRI and other brain science technologies can shed light on how a person's experiences affect cognitive functions that are outwardly undetectable or produce visible symptoms only after it is already too late for a plaintiff to be made whole.\(^6\) In tort law specifically, it is generally more feasible for plaintiffs to recover damages for physical injuries than for psychological or invisible

---


59. It is well-documented that lower-income and racial minority communities, particularly in urban areas, experience an elevated risk for health issues that stem from invisible or late-manifesting harms such as neurotoxins. This Note argues that entrenched policies such as the physical-mental divide in tort law perpetuate these health inequities, since individuals in these communities are also often less able to acquire adequate data to prove the injuries they are suffering from and, moreover, tend to underreport their pain. Emily A. Benfer, *Contaminated Childhood: How the United States Failed to Prevent the Chronic Lead Poisoning of Low-Income Children and Communities of Color*, 41 HARV. ENVTL. L. REV. 493, 503–04 (2017); Jana Mossey, *Defining racial and ethnic disparities in pain management*, 469 CLINICAL ORTHOPAEDICS & RELATED RES. 1859, 1859 (2011).


61. Neuroimaging evidence provides experts with scientific facts upon which they can draw inferences “that not only support the [litigant’s] story but may be the only source for it.” Excluding such evidence would ‘deprive the [litigant] of the voice the Constitution guarantees.’” Adam Teitcher wrote this about criminal defendants but it holds true for civil litigants as well. Adam Teitcher, Note, *Weaving Functional Brain Imaging into the Tapestry of Evidence: A Case for Functional Neuroimaging in Federal Criminal Courts*, 80 FORDHAM L. REV. 355, 393 (2011) (footnote omitted) (quoting CHRISTOPHER SLOBOGIN, PROVING THE UNPROVABLE: THE ROLE OF LAW, SCIENCE, AND SPECULATION IN ADJUDICATING CULPABILITY AND DANGEROUSNESS 55 (2007)).

62. For example, neurocognitive impairment as a result of concussions or neurotoxins may not outwardly manifest until days, weeks, or years later. *See infra* Part II.
injuries, with limited exceptions.\textsuperscript{63} However, with the advent of neuroscientific data, more courts and lawmakers are beginning to see the physical mechanisms underlying pain and humans’ experience of it, expanding plaintiffs’ ability to seek redress for genuine injuries through tort law.\textsuperscript{64} Though the experience of pain and emotion is inherently subjective, this technology can help expose and quantify harms in ways we have not seen before. Furthermore, preventing the worsening of an injury by detecting it at its earlier stages may reduce individual medical treatment costs and curb future litigation.

The next part of this Note will argue why there is inherent value in using this technology to help litigants demonstrate their invisible injuries and finally receive legal recognition.

II. NEUROLAW’S POTENTIAL TO ADVANCE HUMAN RIGHTS IN TORT CASES

This Part describes how neuroscience technology can be used to advance human and civil rights, particularly in the context of tort law. Section A describes the current state of tort standards, and Section B describes how neuroscience is already being presented in courtrooms regarding detection of pain. Sections C, D, and E then explain how neuroscience technology may be useful in helping plaintiffs receive compensation for other invisible injuries, focusing on case studies of emotional distress and PTSD, concussions, and neurotoxicity.

A. Tort Law and the Decline of Substance Dualism

Tort law serves a variety of purposes in society, providing compensation, insurance, and deterrence through an economic model in which those who have caused harm monetarily compensate those

\textsuperscript{63} Two exceptions to the physical-emotional distinction are intentional infliction of emotional distress and negligent infliction of emotional distress, which apply only in rare circumstances, although their allowance has expanded over time and varies by jurisdiction. See \textsc{Restatement (Third) of Torts: Liability for Physical \& Emotional Harm} §§ 46, 47 (Am. Law Inst. 2012).

\textsuperscript{64} “I am confident that we will soon be able to predict, with a high degree of accuracy, some neurological and mental illnesses. Then we will have to answer the question, ‘What do we do now?’” Henry T. Greely, Keynote Address, \textit{Law and the Revolution in Neuroscience: An Early Look at the Field}, 42 Akron L. Rev. 687, 691 (2009).
who have been injured. But what counts as "harm" is not always straightforward. In the United States, "physical" harm has traditionally been distinguished from "mental" harm in tort suits, and courts are much more likely to award damages for physical injuries than for emotional and invisible injuries, largely because the latter two have traditionally been difficult to prove. Relatedly, courts are worried about malingering litigants and are fearful that allowing compensation for less outwardly verifiable injuries could bring a flood of cases that would absorb "resources better left available to those more seriously harmed." Whether it is in statutory law generated by legislatures, common law interpreted by courts, or insurance contracts agreed to between private parties, the law consistently makes this physical-mental distinction.

The theory that mental experiences are something wholly different from bodily ones is known as "substance dualism." Implicit in substance dualism is a societal belief that claims of emotional or mental harm are less deserving of compensation than those with physical consequences. International institutions and American

65. Courts grant compensatory damages for things such as lost wages, medical costs, and loss of earning potential to restore an injured party to his preinjury position. Tort compensation may also "serve an expressive or symbolic function, demonstrating that harming others is a wrongful act that causes dignitary harm beyond the physical and emotional damages." Rick Swedloff & Peter H. Huang, Tort Damages and the New Science of Happiness, 85 IND. L.J. 553, 588 (2010).


68. Comments to the Restatement of the Law (Third) of Torts explain the reasoning behind the distinction: "emotional distress is less objectively verifiable than physical harm and therefore easier for an individual to feign, to exaggerate or to engage in self deception about the existence or extent of the harm." Advances in neuroscience may call this distinction into question as harms such as emotional distress are shown to have a physiological basis. RESTATEMENT (THIRD) OF TORTS (AM. LAW INST. 2012).

69. Notably, the Office of the United Nations High Commissioner for Human Rights recognizes the obligation of states "to protect against [emotional] harm by third parties, including the private sector. . . ." See Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health § 41, U.N. Doc. A/HRC/35/21 (March 28, 2017); see also
courts, to varying degrees, have recognized that tortfeasors should be held responsible for causing emotional distress injuries such as anxiety, loss of tranquility, loss of autonomy, and diminished enjoyment, but have treated these proffered injuries as "second class citizens." Yet nearly all brain researchers and philosophers reject substance dualism in favor of monism, the view that "conscious experience is inseparable from the physical brain." Insights from neuroscience, psychology, and psychiatry have exposed dualism as "empirically flawed and conceptually bankrupt . . . a person cannot be reduced to his mind or separated from his body. He is, inescapably, both at once."

The physical-mental bifurcation in tort law does not sufficiently take into account modern developments in the understanding, diagnosis, and verification of illnesses and trauma that affect brain structure and function. This distinction assumes that medical professionals cannot identify the underlying biological bases of most psychiatric disorders, as opposed to the more discernible pain of an outwardly obvious injury like a broken leg. As we continue to discover the physiological origins of emotional harm through

Lisa J. Laplante, Human Torts, 39 CARDOZO L. REV. 245, 247 (2017) (arguing that emotional distress suits can "be reframed as violations of the most basic human rights such as the right to physical and mental integrity under international human rights law, grounded in treaty and international customary law").

70. Grey, supra note 58, at 2605–08.

71. JAMES W. KALAT, INTRODUCTION TO PSYCHOLOGY 6 (9th ed. 2010). As a "learned author" quoted by Judge Clark in Young v. W. Union Tel. Co. puts it, "the mind is no less a part of the person than the body, and the sufferings of the former are sometimes more acute and lasting than those of the latter." 107 N.C. 370, 385, 11 S.E. 1044, 1048 (1890).

72. Dov Fox & Alex Stein, Dualism and Doctrine, 90 IND. L.J. 975, 975–1010 (2015).

73. Scientists and scholars are moving away from treating "mental" and "physical" as separate categories. See Peter A. Alces, THE MORAL CONFLICT OF LAW AND NEUROSCIENCE 131–32 (University of Chicago ed., 2018) ("Once we have a way to 'see' emotional injury as clearly as we can 'see' a broken bone . . . there would be no reason to maintain the tort law's distinction between physical and emotional injury."); see also Govind Persad, Law, Science, and the Injured Mind, 67 ALA. L. REV. 1179, 1215–16 (2016) ("Our improved understanding of the biological correlates of mind-dependent harms suggests that the line between 'body' and 'mind' is no longer sufficient to support the differential legal treatment of these harms.").
brain imaging evidence, the distinction becomes outmoded. As advancements in neuroscience grant us greater ability to quantify emotional harm claims, tort litigants may be able to introduce neuroimaging evidence as objective proof of injury, and courts and legislatures should begin to see emotional injuries’ capacity to harm litigants and rethink this anachronistic distinction. Many of these litigants are members of populations that are underprivileged or that have historically been undercompensated in their legal rights. If neuroscience technology continues to be increasingly precise and reliable, then many litigants whom have suffered trauma that has


75. See Betsy Grey, *Implications of Neuroscience Advances in Tort Law: A General Overview*, 12 IND. HEALTH L. REV. 671, 689–90 (2015) (arguing that distinctions between emotional and physical pain are false because of the changes in the brain that result from emotional pain). The court’s analysis in Allen v. Bloomfield Hills School District, 760 N.W.2d 811 (Mich. Ct. App. 2008), may signal the beginning in rethinking the physical-emotional divide. Allen, diagnosed with PTSD stemming from an accident in which he suffered no bodily injuries, submitted a PET scan of his brain depicting abnormal decreases in frontal and subcortical activity which the Court of Appeals said represented “objective medical evidence that a mental or emotional trauma can indeed result in physical changes to the brain” and found that “[t]he brain is a part of the human body, so ‘harm or damage done or sustained’ is injury to the brain and within the common meaning of ‘bodily injury’ in MCL 691.1405 ... What matters for a legal analysis is the existence of a manifest, objectively measured injury to the brain.” Allen, 760 N.W.2d at 815. See also Pekin Ins. Co. v. Hugh, 501 N.W.2d 508, 512 (Iowa 1993) (finding that whether a claimant suffered “bodily injury” involved “a medical or psychological problem of proof rather than purely a question of law ... [Compensation] should not therefore turn on any artificial and arbitrary classification such as ‘physical’ or ‘psychological’”). The distinction is also losing traction in the international sphere: The Supreme Court of New South Wales ruled that a woman who experienced PTSD as the result of an airplane crash could recover damages under the Montreal Convention because it deemed PTSD was in and of itself “bodily injury,” unlike previous rulings under the international agreement. Victoria Gallanders, *Australia: Post Traumatic Stress Disorder (PTSD) ruled as bodily injury in landmark case*, MONDAQ (July 7, 2015), http://www.mondaq.com/australia/410418/Personal+Injury/Post+Traumatic+Stress+Disorder+PTSD+ruled+as+bodily+injury+in+landmark+case [https://perma.cc/N3N6-VGTW].

76. AM. PSYCHIATRIC ASS’N, *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS* (5th ed. 2013) (incorporating a broadened view of mental injury as a result of these neuroscientific advances).

77. For example, neuroscience can help inform administrators about the emotional impact of solitary confinement on prisoners and the emotional effect of prison violence, which could lead to a more rehabilitative and effective criminal justice system. See Gertner, supra note 10, at 544–46.
otherwise been overlooked by the legal system may at least be able to get their day in court.

Indeed, neuroscience research has begun to shed light on the specific neural correlates of emotional pain, proffering concrete evidence that these injuries may not be so second-class. Individuals with emotional trauma or psychiatric disorders have abnormalities in a number of brain regions, including the adrenal systems, the amygdala, the hippocampus, and the cortices, and exposure to traumatic events can change this circuitry in previously healthy individuals. Studies have confirmed that, in certain instances, pugnacious words or tones, verbal threats, bullying, or emotional abuse can cause neurochemical changes in the amygdala and atrophy in prefrontal cortical function. This can influence students' ability to perform in school and affect an individual's likelihood of future disease or even one's lifespan. fMRI scans show that emotional pain physically affects the same brain area as bodily injuries and that emotional harm can be at least as painful as physical harm. Emotional pain can also be more prolonged or more debilitating than physical pain.

As diagnostic imaging techniques are increasingly able to provide more objective evidence of these kinds of brain-based distress, there is less justification for entrenched laws and policy that make it harder to recover for invisible injuries than physical ones. Limiting
tort claims to outwardly visible injuries fails to recognize the progress that our society has made in understanding mental health matters and that non-visible injuries can be as real and debilitating as visible ones. 84 How neuroscientific evidence is used in courtrooms and translated into policy may thus have significant ramifications for civil recovery, especially for those litigants who have been effectively disparaged because they had no objective evidence to substantiate their claims. 85 At the very least, the increased ability to detect and quantify emotional pain should force courts and legislators to reexamine this divide. The following are examples of ways that neuroscience can help detect evidence of invisible injuries and preserve the deterrent and corrective justice functions of civil law, particularly in the field of torts.

B. Neuroscience and the Identification of Pain

One area of civil litigation in which neuroscience increasingly plays a role is in the determination and valuation of pain. About $150 billion 86 and hundreds of thousands of legal proceedings each year 87 in

assessments of physical pain, emotional distress, and a variety of psychiatric disorders” that are largely subjective experiences). 84 AM. PSYCHIATRIC ASS’N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS xxx (4th ed. text rev. 2000) (stating “the term mental disorder unfortunately implies a distinction between ‘mental’ disorders and ‘physical’ disorders that is a reductionistic anachronism of mind/body dualism. A compelling literature documents that there is much ‘physical’ in ‘mental’ disorders and much ‘mental’ in ‘physical’ disorders”). Judge Totenberg referenced this passage in her decision in Reid v. Metro. Life Ins. Co., 944 F. Supp. 2d 1279, 1305 (N.D. Ga. 2013).

85 Mark Anderson, who served as a mental health policy advisor in the Senate, stated that “a health care system that does not treat the brain with the body is outmoded.” Francis X. Shen, Mind, Body, and the Criminal Law, 97 MINN. L. REV. 2036, 2061 (2013). This has gained some traction at least in the health law sphere as there has been a sustained movement to enact mental health parity laws that recognize “biologically based mental illness.” Id. at 2060; see also Reid, 944 F. Supp. 2d at 1323, in which a plaintiff used neuroscientific data to demonstrate that her diagnosis of dementia is based in physical processes in order to receive long-term disability benefits. Judge Totenberg placed great weight on an MRI of Ms. Reid’s brain, stating that Ms. Reid’s dementia was “confirmed by her neuroimaging results showing cerebral atrophy,” and that MetLife’s failure to consider the MRI as evidence of dementia was arbitrary and capricious. Id.


87 Adam Kolber estimates that pain is an issue in about half of all tort cases. Greg Miller, Brain Scans of Pain Raise Questions for the Law, 323 SCI. 195, 195 (2009).
the United States turn on the existence and extent of a litigant's pain, often in disability, insurance, product liability, medical malpractice, workers' compensation, personal injury, and pain and suffering proceedings. Yet, evaluation of pain largely depends on subjective self-reporting that can frequently be difficult to express, relatively easy to exaggerate, and hard for others to understand.

Neuroscience may be able to corroborate a person's described experience of pain or provide evidence about pain for those who are unable to verbally communicate, such as infants or the unconscious. The use of brain imaging technology can demonstrate where in the brain litigants' self-reported pain correlates with neural activity and how particular pain conditions result in the reshaping of certain brain structures and neural circuitry. Researchers have found, for example, that chronic headaches, back pain, and phantom limb pain are associated with decreased grey matter density in the prefrontal cortex and thalamus of the brain. Scholars have advocated that neuroimaging of chronic pain should lead to modifications of mental

88. Amanda C. Pustilnik, Pain as Fact and Heuristic: How Pain Neuroimaging Illuminates Moral Dimensions of Law, 97 CORNELL L. REV. 801, 801 (2012) (“Important legal distinctions turn on the presence and degree of physical pain. [For example, some] statutes refer to degrees of physical pain to define criminal offenses like torture-murder, while pain that rises to the level of cruelty draws the boundary between constitutionally permissible and impermissible punishment.”).

89. “Pain and suffering” includes fright, nervousness, grief, anxiety, worry, mortification, shock, humiliation, indignity, embarrassment, apprehension, terror or loss of enjoyment of life that a tort victim suffers because of the civil wrongdoing of another. Adam J. Kolber, Pain Detection and the Privacy of Subjective Experience, 33 AM. J.L. & MED. 433, 441 (2007).

90. “[P]lain is largely invisible, unquantifiable, and often grossly misunderstood, leading to unnecessary suffering on the part of people whose pain is not credited and to unnecessary expense when the legal and medical systems function inefficiently or the wrong claimants are compensated.” Amanda C. Pustilnik, Painful Disparities, Painful Realities 3 (U. Md. Legal Stud. Res. Paper No. 2014-18, 2014).

91. Id. at 4.

92. A. Vania Apkarian et al., Chronic Back Pain Is Associated with Decreased Pre-frontal and Thalamic Gray Matter Density, 24 J. NEUROSCI. 10410, 10412 (2004). PET, electroencephalography, and fMRI scans have shown that the cortical and subcortical regions, regions responsible for handling sensory perceptions, activate during pain stimulation and are referred to as the “pain matrix.” Parts of these regions can discern the location and intensity of painful stimuli while others are involved in the experiential and perceiving aspect of pain. Tor D. Wager et al., An fMRI-Based Neurologic Signature of Physical Pain, 368 NEW ENG. J. MED. 1388, 1388 (2013).
and physical disability regulations as well as workers’ compensation regimes.93

Although technology has allowed researchers to discover structures of the brain that are responsible for pain perception, whether this is enough to prove pain and suffering in court is more tenuous, especially under the Daubert and Frye standards94 for the admission of expert evidence.95 Expert opinions diverge on whether neuroimaging technology’s ability to determine precise levels of pain is reliable enough for courtroom use. Some argue that many external factors affect an individual’s pain perception96 and that sensitivity to pain varies significantly from one individual to another. Professor Amanda Pustilnik believes that aggregate pain neuroimaging evidence ought to be admissible under the federal, state, and administrative evidence regimes for limited purposes, but that brain scanning technology is not, or at least not yet, a “fraud-o-meter, pain-o-meter, or mind-reading machine;” it is better used as a tool for increasing understanding about these complex phenomena and for educating judges and jurors.97 A number of private entities,98 scientists,99 and

93. See generally Kolber, supra note 83, at 587 (noting that technological advances in neuroscience will improve assessments of physical pain, emotional distress, and psychiatric disorders that are subjective experiences) and Tovino, supra note 44, 697–727 (exploring role that brain scanning technologies play in health insurance coverage, social security eligibility and recognizing gender-specific mental health conditions).

94. See supra note 30 (defining the Daubert and Frye standards).

95. It is also important to note that a majority of pain-related claims are heard in administrative settings for matters such as workers’ compensation and disability claims, in which the “rules of evidence are slacker, and in some cases close to non-existent, and there is no jury, and there are no instructions.” David Seminowicz et al., Panel 1: Legal and Neuroscientific Perspectives on Chronic Pain, 18 J. HEALTH CARE L. & POLY 207, 225 (2015). Administrative law judges are not bound by Daubert and can be crucial decision makers in applying the law relative to evidence of chronic pain. Id. at 226.

96. Psychological factors including anxiety, attention, and distractions may alter signals in fMRIs. Miller, supra note 87, at 195.


98. A number of private companies, such as Connecticut-based company Millennium Magnetic Technologies, offer brain scanning services for litigants to validate the presence of pain. See Use of Functional MRI to Validate the Presence of Pain, MMT NEUROTECH, www.milmag.net/document-pain [https://perma.cc/6VSP-JUVL].

99. Studies by neuroscientists such as Tor Wager at the University of Colorado and Sean Mackey at Stanford University have also determined, at least in controlled experiments, that fMRIs were able to determine with 80% accuracy whether or not a particular subject is in pain. Sara Reardon, Neuroscience in Court: The Painful Truth, 518 NATURE 474, 475 (2015).
courts have taken more lenient views, and the use of pain-scanning techniques for tort litigation has risen.

fMRI pain scans may not meet the requisite level of certainty to make them useful in all cases. Even so, the reliability of the technology has increased markedly in a very short period of time. While today's neuroscience has not yet produced a foolproof "pain-o-meter," it can map brain pathways and offer increasingly tailored treatment, explanation, and measurement of pain in individuals. It "now shows that distinct chronic pain conditions produce characteristic patterns of structural brain alteration, with the degree of visible brain alteration correlating with the duration, severity, and type of chronic pain," which can offer some visibility to litigants' previously discounted claims of pain.

C. Neuroscience and PTSD

Cases involving PTSD illustrate how broken the physical-mental division in tort law is. Neuroscience research has begun to document structural changes to the brains of PTSD-diagnosed participants by showing how PTSD results from disrupted circuitry between the amygdala, the brain region stimulated when an individual experiences stress or trauma, the hippocampus, which plays a

100. Carl Koch, whose wrist was burned by molten asphalt, sued his former employer for damages for his chronic pain over a year after the burn. Judge Chon-Lopez admitted Koch's brain scan indicating his pain. Id. The case ultimately settled for $800,000. Id.

101. Id.

102. Karen D. Davis et al., Brain Imaging Tests for Chronic Pain: Medical, Legal and Ethical Issues and Recommendations, 13 NATURE REV. NEUROLOGY 624, 634 (2017) (asserting that "[neuroscience] research can guide the crafting of more accurate and precise laws that relate to pain as a source of disability, and can assist the evaluation of evidence in individual cases").

103. Pustilnik, supra note 30, at 1117.

104. Since its official recognition in 1980 by the American Psychiatric Association, PTSD has become prevalent in much personal injury litigation. From 1999 to 2004, there was nearly an 80% increase in PTSD cases, with payments reaching over $4 billion, and those numbers have continued to expand. Harvard professor Alan Stone has remarked that "no diagnosis in the history of American psychiatry has had a more dramatic and pervasive impact on law and social justice than . . . PTSD." Shen, supra note 85, at 2159 (footnote omitted).

105. Research has shown that stress impairs the hippocampus's capacity to regenerate neurons as part of its normal functioning. Studies of children with PTSD in particular have found that these impairments can lead to problems with learning, memory, and academic achievement. CHILD WELFARE, UNDERSTANDING THE EFFECTS OF MALTREATMENT ON BRAIN DEVELOPMENT 8–9 (2015),
central role in the formation of memory, and the pre-frontal cortex, which regulates emotional responses to fear and stress. Neuroscientists have found that PTSD can also cause disruption of neurotransmitter networks such as the noradrenergic system, the serotonergic system, and the hypothalamic-pituitary-adrenal axis.¹⁰⁶

Many types of psychological trauma can cause PTSD, such as car accidents, military combat, childhood abuse, rape, and assault.¹⁰⁷ Patients with PTSD can suffer from a wide array of symptoms that reflect stress-induced changes in neurobiological systems, including "intrusive memories, flashbacks, hyper-vigilance, sleep disturbance, avoidance of traumatic stimuli, physiological hyperresponsivity, numbing of emotions, and social dysfunction."¹⁰⁸ While such symptoms are commonly understood to be psychological problems, some or all of them may well be related to the physical effects of extreme stress on the brain resulting from a dysfunction of the neural networks that regulate memory and fear.

Neuroscience evidence can play a critical role in determining whether PTSD is understood as a mental or a bodily injury, the latter of which is often required for plaintiffs to recover damages in tort claims as well as in insurance policies, contracts, or claims against the government. To receive damage awards in PTSD litigation, a plaintiff must establish the existence of PTSD as well as specific causation between the defendant's actions and that injury. Neuroscience findings can help fortify that chain.¹⁰⁹ Documenting real-time neurochemical changes can help clarify that there is at least some physical damage that correlates with PTSD in an individual. These advances might be particularly salient for victims of domestic violence as medical professionals are becoming increasingly cognizant of the neurobiological consequences of battering.¹¹⁰ Furthermore, while


¹⁰⁶. PTSD "causes significant changes in brain chemistry, brain function, and brain structure. The brain becomes 'rewired' to over-respond to circumstances that are similar to the traumatic experience." Allen v. Bloomfield Hills Sch. Dist., 760 N.W.2d 815, 816 (Mich. Ct. App. 2008) (footnote omitted).


¹⁰⁸. Id.

¹⁰⁹. Shen, supra note 66, at 332.

society might consider military veterans or victims of domestic violence to be "deserving victims" of compensation for PTSD, what about victims of PTSD triggered by their imprisonment or participation in gang violence? Should social disapproval of these stressors allow the criminal justice system to withhold sentencing mitigation of these defendants even though they too suffer from PTSD? Neuroscience can offer credence to the claims of such victims and to people of all backgrounds, regardless of their status in society.

This increasingly substantial research suggests that maybe it is time to stop disfavoring tort liability for emotional distress, or at least reconsider it for particularly vulnerable populations. The capacity of functional neuroimaging to identify the biological correlates of emotional experience makes it clear that being subjected to stress or abuse can lead to objectively identifiable changes in the brain and can play a role in substantiating subjectively reported pain. We must critically examine and clarify the normative foundations for the distinctions we have historically taken for granted.

111. This issue has come up in international settings as well. In a 1998 trial of a Bosnian-Croatian soldier, experts for both the defense and the prosecution relied on neuroscientific evidence to argue whether a torture victim was suffering from PTSD. Shen, supra note 66, at 333.

112. Vulnerable populations might include those that have traditionally had limited access to courts or have been relatively neglected; they might include persons of lower economic backgrounds, racial, gender, or ethnic minorities, or those in overlooked areas of society. Prisoners, for example, are potentially given less credence in courtrooms than other citizens: the Prison Litigation Reform Act prohibits prisoners from suing for emotional injury without being able to show physical injury or sexual misconduct. 42 U.S.C. § 1997e (2012). See also Persad, supra note 73, at 1199 (discussing how U.S. courts tax the damages victims of emotional injury receive, while leaving damages for physical injury untaxed).

113. Betsy Grey, Neuroscience and Emotional Harm in Tort Law: Rethinking the American Approach to Free-Standing Emotional Distress Claims, in 13 LAW & NEUROSCIENCE: CURRENT LEGAL ISSUES 203, 225 (Michael Freeman ed., 2011) (arguing that "the availability of neuroimaging evidence should argue in favor of abandoning the more artificial and arbitrary tests for limiting emotional harm claims such as physical impact, physical manifestation and zone of danger").

might be necessary, the physical-mental distinction may no longer be the most appropriate point on which to divide it.

D. Neuroscience's Potential to Identify Mild Traumatic Brain Injury

Litigation and national public health concerns over mild Traumatic Brain Injury (mTBI) and Traumatic Brain Injury (TBI) have gained prominence as reliance on brain scanning has increased. The terms "concussion" and "mTBI" are often used interchangeably. mTBI currently stands as a subjective clinical diagnosis based primarily on patient history and observable behavioral symptoms, which may include concussions and loss of consciousness, confusion,

---


dizziness, fatigue, nausea, and trouble with learning and memory.\textsuperscript{117} Sometimes symptoms might endure for weeks or longer, manifesting in persistent headaches, sleep disturbance, poor concentration, irritability, and depression.\textsuperscript{118} The Centers for Disease Control and Prevention estimates that, in the United States, 1.7 million people suffer a TBI each year, and it is likely that many more go unreported.\textsuperscript{119} With an estimated total cost to society exceeding $76 billion per year, the epidemic has far-reaching consequences.\textsuperscript{120}

Unfortunately, because mTBI manifests in no particular physically distinct way, there is presently a lack of unequivocal metrics to detect it.\textsuperscript{121} Thus, the millions of professional and youth athletes, as well as victims of head trauma in transportation accidents, military combat,\textsuperscript{122} domestic abuse, and workplace injuries, can find tort claims difficult to prove in court.\textsuperscript{123} This is especially so when debilitating symptoms, which can be extremely severe in some cases,\textsuperscript{124} do not

\begin{itemize}
  \item \textsuperscript{117} Richard P. Dutton et al., \textit{Diagnosing Mild Traumatic Brain Injury: Where Are We Now?}, 70 J. TRAUMA 554, 554 (2011).
  \item \textsuperscript{118} \textit{Id.}
  \item \textsuperscript{119} CTRS. FOR DISEASE CONTROL & PREVENTION, GET THE STATS ON TRAUMATIC BRAIN INJURY IN THE UNITED STATES, https://www.cdc.gov/traumaticbraininjury/pdf/bluebook_factsheet-a.pdf [https://perma.cc/2YHD-7HYC].
  \item \textsuperscript{120} \textit{Severe TBI}, CTRS. FOR DISEASE CONTROL & PREVENTION, http://www.cdc.gov/TraumaticBrainInjury/severe.html [https://perma.cc/6NWB-9774].
  \item \textsuperscript{121} "[m]TBIs often do not show up on . . . CT . . . or MRI [scans], since the injuries are typically not structural injuries to the brain, but rather, are functional problems caused by swelling or bruising." Grey & Marchant, \textit{supra} note 116, at 1924. Neuroscientist Michael Selzer has said that "cognitive deficits can be subtle, even to a neurologist." Emily Singer, \textit{Detecting Subtle Brain Injuries}, MIT TECH. REV. (Nov. 18, 2008), https://www.technologyreview.com/s/411201/detecting-subtle-brain-injuries/ [https://perma.cc/LMX7-PGKJ].
  \item \textsuperscript{123} See Boyd v. Bert Bell/Pete Rozelle NFL Players Ret. Plan, 410 F.3d 1173, 1177 (9th Cir. 2005) (noting that Boyd's SPECT scan revealed decreased brain activity consistent with head trauma and was admitted into court under \textit{Daubert}, but the physician for the defense claimed that, based on this evidence, the head injury in question "could not be organically responsible for all or even a major portion of the neurologic and/or neuropsychologic problems that Mr. Boyd is experiencing now, to a reasonable degree of medical probability"). As neuroscience imaging becomes more precise, however, this type of evidence may help plaintiffs with similar injuries in the future.
  \item \textsuperscript{124} The estate of former NFL player Aaron Hernandez filed a lawsuit against the New England Patriots and the NFL, claiming that Hernandez's injuries and death were a direct result of his participation in football and that the Patriots
manifest immediately. For those that suffer from chronic traumatic encephalopathy (CTE), their diagnosis and resulting redress may come too late.\footnote{Symptoms of CTE, which include memory loss, depression, suicidal thoughts, and aggressive behavior, have been noted in ice hockey players, soccer players, boxers, and football players, among others. CTE symptoms in the brain reflect those found in Alzheimer’s patients and can take years or even decades after the brain trauma has occurred to manifest. CTE currently can be diagnosed definitively only through autopsy. Grey & Marchant, supra note 116, at 1914–15, 1918.}

Over the past few years, however, neuroscientific research has begun to develop more nuanced, objective diagnostic measures\footnote{Id. at 1936. Researchers used PET scans after injecting a radioactive tracer that binds to deposits of tau, a protein secreted by the axons of unmyelinated nerve cells when they are injured. Using these PET scans, the researchers were able to pinpoint where in the brain these abnormal proteins accumulated and identify distinctive patterns of tau tangles in the amygdala and subcortical regions of the former football players that did not appear in the normal brains of the study’s controls. Other studies have used S100B and glial fibrillary acidic protein and found that increased levels of SNTF, a protein that increases in the blood after some concussions, were strongly correlated with diffuse axonal injury and long-term cognitive dysfunction.} of mTBIs by, for instance, evaluating levels of proteins that regulate cell development and degradation and act as an indicator of TBI.\footnote{These include MRIs, diffusion tensor imaging, magnetization transfer ratio, susceptibility weighted imaging, myelin water imaging, ultrashort echo time, and proton magnetic resonance spectroscopy that detects white matter injury. Ponnada A Narayana, White matter changes in patients with mild traumatic brain injury: MRI perspective, 2 CONCUSSION 2 (2017).}

Neuroscientists also scan eye movements of patients to detect signs of mTBI via magnetoencephalography, a functional brain imaging technique that measures neuronal currents. A variety of other neuroimaging techniques are also being investigated for providing biomarkers of mTBI.\footnote{Grey & Marchant, supra note 116, at 1937.} In the fall of 2017, a group of researchers at Boston University discovered a possible means of detecting CTE in
living plaintiffs; 129 studies are increasingly finding that diffusion tensor imaging may also be useful in detecting mTBI. 130 This is one example of an advancement in neuroscience enabling a far larger class of plaintiffs to litigate. These methods are in their infancy, however, and more studies are needed before they can be used as definitive ways to diagnose mTBI, TBI, and CTE in living patients.

As the science improves, more precise brain scans may significantly inform courts and policymakers as they wrestle with complex questions regarding the nature of concussive injury, the need for regulation in the area, and the allocation of fault and duties with regard to head injuries. Indeed, all fifty states now have legislation to prevent concussions and to limit further injury to student-athletes who sustain concussions. 131 Moreover, with increasingly accurate brain scanning technology, more claims may be made in professional malpractice lawsuits and in negligence claims against entities sponsoring athletic events. Plaintiffs may also pursue latent injury claims, 132 which have been causally difficult to substantiate. These developments might also prompt courts to hold coaches, trainers, and parents to a higher standard of care. The findings will help scientists better define the type of damage that can lead to long-lasting memory and emotional problems, as well as help identify those who are most vulnerable to further trauma.

Ultimately, neuroscience will inform the debate as to what risks are acceptable for whom, and it will allow individuals to better understand how their brains have been affected. It will also prompt more just compensation, especially considering the coercive social and economic pressures to underreport symptoms or to ignore the risks of


132. "Latent injury claims permit a claimant to receive compensation before a serious disease has manifested." Id. at 1958.
these activities, which are often inevitably linked to issues of social and racial justice.

E. Neuroscience’s Potential to Identify Neurotoxins in the Brain and Nervous System

Exposure to toxins can cause neurophysiological changes in the brain, which can lead to cognitive impairment, neurodegenerative diseases, or the onset of psychiatric disorders such as ADHD or autism—even though these effects are typically invisible to the naked eye. Common pesticides and chemicals such as lead, arsenic, organophosphates, and mercury have been closely associated with cognitive impairment, and more than 200 chemicals have been shown to be neurotoxic in humans. It is not just factory pollutants that can

133. Some athletes, for example, feel they have no other choice but to sacrifice their bodies and minds to make money or have a future, if they are even aware of the risks in the first place. See Jesse Dougherty, Former Alabama player Les Williams is one of more than 100 suing NCAA over brain injuries, WASH. POST (July 2, 2018), https://www.washingtonpost.com/news/sports/wp/2018/07/02/feature/former-alabama-player-les-williams-is-one-of-more-than-100-suing-ncaa-over-brain-injuries/?utm_term=.b8960128a6e9 [https://perma.cc/6DFG-BGTV].

134. Alana Semuels outlines where many of the fault lines occur, particularly in football and other high-contact sports, for those who have the opportunities to avoid the grave repercussions playing football might bring, and those who do not: “the divide on the football field makes it hard not to see how inequality in America is worsening health disparities and raising the specter of another, darker era of American history.” Alana Semuels, The White Flight from Football, ATLANTIC (Feb. 1, 2019), https://www.theatlantic.com/health/archive/2019/02/football-white-flight-racial-divide/581623/ [https://perma.cc/M2BG-FMNU].


136. “Chronic exposure to lead has measurable effects on the nervous system due to lead’s propensity to accumulate in bone over time. For instance, in an MRI study of 532 former lead workers, high tibia lead was associated with reduced total brain volume, lower volume of gray matter in the insula and cingulum, and diminished white matter volume in the parietal lobes.” Despite the fact that lead levels have been largely reduced in the United States, thanks in great deal to better science detailing its deleterious effect on brain development, the presence of lead in water, soil, and wall paint continues to affect many. Lisa H. Mason et al., Pb Neurotoxicity: Neuropsychological Effects of Lead Toxicity, BIOMED RES. INT’L, Jan. 2, 2014, at 2.

137. Gennaro Giordano & Lucio G. Costa, Developmental Neurotoxicity: Some Old and New Issues, 12 ISRN TOXICOLOGY 1, 3 (2012). Exposure to subclinical levels of toxins such as manganese, methylmercury, polychlorinated biphenyls, ethanol, lead, arsenic, toluene, fluoride, chlorpyrifos, and tetrachloroethylene has been shown to disrupt brain development and normal neurotransmitter function.
have deleterious effects on the brain—chemicals that are found in children’s toys, in food, and even in household furniture and appliances can also affect neural development. 138 “[T]he vast majority of chemicals in commerce remain untested for their impacts on neurodevelopment,” 139 and it is quite possible that there are more neurotoxins whose effects have not yet been identified.

The prevalence of neurotoxins and their often-deleterious effects on the body and mind can elude public, and certainly legal, discourse. In fact, the magnitude of neurotoxicity is not exactly known and its impact on human health can be understated; even in more developed countries, it is estimated that over 30 million individuals suffer from neurobehavioural illness, but only 20% of these individuals seek medical attention related to such illness. 140 Because many neurotoxins cause non-distinct clinical manifestations—such as nausea, headaches, pain, irritability, dizziness, fatigue, and difficulty concentrating—it is often difficult to diagnose with reasonable certainty, as is required in tort law to recover damages, whether someone is suffering from exposure to toxins. Aside from issues related to proving causation, toxic tort litigation can also require extensive preparation and testing, substantial financial resources, and expert testimony to interpret the evidence, which can be prohibitively expensive for many litigants.

As reliance on neuroscience technology and neurobiological research gain is augmented, more litigants will be able to overcome these obstacles and bring suit, 141 which may in turn affect legislation


138. Bisphenol A, a ubiquitous chemical found in many plastics, has neurobiological effects. Linda S. Birnbaum et al., Environmental Health Science For Regulatory Decisionmaking, 21 DUKE ENVTL. L. & POLY F. 259, 279 (2011).

139. Id.


141. “It is a given that many subclinical events, once considered invisible and thus speculative, will become detectable and hence objectively verifiable.” Jamie A. Grodsky, Genomics and Toxic Torts: Dismantling the Risk-Injury Divide, 59 STAN. L. REV. 1671, 1704 (2007). Grodsky wrote this about genetic testing, but it remains true for neuroscientific testing as well. See, e.g., Hose v. Chi. Nw. Transp. Co., 70 F.3d 968, 973 (8th Cir. 1995) (finding no abuse of discretion where the district court admitted PET evidence to show injuries consistent with manganese encephalopathy); In re Welding Fume Prods. Liab. Litig., 245 F.R.D. 279, 298 n.111 (N.D. Ohio 2007) (considering MRI and PET scans demonstrating whether
or discourse about the acceptable levels of toxic chemicals that can be discharged into the environment. Chronic levels of exposure to toxins such as pesticides, even at low levels, can have profound impacts on the nervous system, especially for children and infants. Health effects may occur years after minor exposure to toxins in the environment or in residues ingested through food and water. Neuroscience studies illustrate dysfunctions in brain and behavior that may be attributable to ecological toxins, and such studies have found that a number of toxins are risk factors for the development of neurodegenerative diseases such as Alzheimer's or Parkinson's later in life. Chronic exposure to neurotoxic substances can also be associated with violence, depression, and substance abuse.

exposure to welding fumes and manganese can cause, contribute to, or accelerate a Parkinsonian syndrome).

142. For example, researchers have used fMRIs to detect the effect of prenatal methylmercury exposure in adolescents and have used MRIs to examine how the chemicals appear to cause thinning of the cortex in children's brain structure. Hamblin, supra note 137.

143. See Laura Y. Cabrera, Pesticides: A Case Domain for Environmental Neuroethics, 26 CAMBRIDGE Q. HEALTHCARE ETHICS 602, 603 (2017) (“Pesticides readily cross the placenta and bioconcentrate in breast milk, resulting in early-life exposure during critical prenatal neurodevelopment. . . . Research indicates that children born to mothers exposed to pesticides during pregnancy . . . [can] lag . . . two years behind in motor and spatial development when compared with children of mothers without [this] exposure. Other studies have found an association between residential proximity to agricultural fields where exposure to pesticides during pregnancy was correlated with autism spectrum disorder.”) (footnotes omitted).

144. In the case of organophosphates, for example, individuals might develop impaired cognitive and psychomotor function a few weeks after exposure, and, in some cases, effects were observed ten or more years after poisoning, suggesting that the residual damage is permanent. Id. at 605.

145. For example, neuroscientists use brain scans to assess the patient’s sensory, motor, reflex and cranial nerve function to diagnose and monitor substances like lead and mercury on the peripheral nervous system. Anetor et al., supra note 140, at 10.

146. Baskin-Sommers & Fonteneau, supra note 135, at 431 (“Research attributes exposure to synthetic chemicals, including those found in drugs and pesticides, to damage of dopaminergic neurons in the nigrostriatal system . . . depletion of dopamine in the SN pars compacta and subsequent cell death . . . . Additionally, beta-amyloid protein plaques and intracellular neurofibrillary tangles are linked to toxic environmental exposure, as is inflammation of the brain and accumulation of trace metal elements in brain regions, such as the basal ganglia.”).

Similar to victims of emotional trauma, PTSD, and mTBIs, neurotoxin victims often face difficulties—whether under common law, contract, or policy—when their injuries are not physically visible and thus have been historically difficult to demonstrate. In cases of asbestosis, for example, insurance coverage is often triggered only when the disease causes "bodily injury" even though the policy does not delineate exactly what counts as such injury and when it occurs. Some courts have ruled that the microscopic tissue damage caused by asbestos exposure constitutes bodily injury, whereas others have ruled the opposite, finding that a bodily injury must be "an injury, sickness, or disease," when one's sense of well-being is adversely affected or impaired, and thus have preferred a principle of severity. The problem with this is that many neurological injuries are insidious, and not severe or compensable until it is too late—that is, their debilitating effects have already begun and may not be curable by the time a litigant can prove their existence in court.

The existence of these toxins in the environment represents an important challenge to environmental justice and human rights. There are gross inequities between resource-poor and industrialized countries as well as between different socioeconomic and racial

148. Daniel A. Farber, Toxic Causation, 71 MINN. L. REV. 1219, 1247 (1987) ("The only real difference between the automobile case and the toxics case is that better information is available about the events in the automobile case whereas the relevant biological events in the toxics case are unobservable.").

149. Shen, supra note 85, at 2118. Asbestos is a prime example of a latent injury claim.


151. This latency is exacerbated by the sluggish, and often ineffective, review process that the Environmental Protection Agency (EPA) and other agencies go through to determine whether to ban a substance or not, such as in the case of the highly toxic insecticide, chlorpyrifos. Megan K. Horton et al., Neuroimaging is a novel tool to understand the impact of environmental chemicals on neurodevelopment, 26.2 CURRENT OPINION IN PEDIATRICS 230, 233–34 (2014).

152. Resource-poor countries often have rudimentary pesticide registration, regulation, handling, and enforcement requirements. Unsurprisingly, they can experience pesticide overload per capita, including from pesticide products that are banned elsewhere. Furthermore, pesticide labels often are not listed in the local language nor are written in a manner easily understandable by average consumers. Cultural and contextual considerations are often overlooked as well when distributing pesticides to these locations. For example, climatic conditions or limited budgets might make it impractical to wear suggested protective clothing. Discrepancies such as these contribute to further disparities in pesticide burden. Cabrera, supra note 143, at 606.
groups within countries. Those with lower socioeconomic status are more likely to live in areas where toxins are present, to have occupations that involve direct contact with toxins, and to have overall less bargaining power and access to education to treat or prevent them. The incidence of lead poisoning, for example, is associated with "socioeconomic status, rurality, race, age, and the date one's residence was built." Poorer urban children are at the highest risk for neurotoxicity, "presumably due to the presence of lead in older building materials and reduced access to sources of nutrition." Exposure to these toxins can further magnify inequalities and hamper educational opportunities, especially considering the young age of many lead poisoning victims.

Although there is substantial evidence demonstrating the impact of certain pesticides on brain and mental health, the synergistic and cumulative effects of many toxins can be difficult to measure and might persist for years before detection. We need better and more preventative methods to determine which chemicals have neurotoxic effects, and neuroscience research and scanning technologies can

153. For example, it is well documented that there are disproportionate levels of lead paint and landfills (from which toxic substances such as mercury can seep) located in communities of color or lower-income populations. People in these communities may already be at a disadvantage in obtaining costly medical evidence and are prone to suffering long-term neurological illnesses. Disparities such as these are but one window into the type of neurological injuries that some communities disproportionately face, and could be remedied by tort law that is more proactive in recognizing invisible injuries. See generally Benfer, supra note 59 (documenting above); Robert D. Bullard, Race and Environmental Justice in the United States, 18 YALE J. INT'L L. 319, 334 (1993) (explaining that low-income and minority communities suffer most from the nation's environmental problems and have not had success in preventing construction of waste and other polluting facilities).

154. Prisoners, and those who live and work near prisons, for example, are particularly vulnerable to toxin exposure. See Prison Ecology Project, NATION INSIDE, https://nationinside.org/campaign/prison-ecology/ [https://perma.cc/C3N7-FKKH].

155. Mason et al., supra note 136, at 3.

156. In the 1970s, the average U.S. preschool child had 15 micrograms of lead per deciliter of blood, eighty-eight percent of children had a level exceeding 10 \( \mu g/dL \)—twice what the CDC currently considers toxic—and the average level was markedly higher at 23 \( \mu g/dL \) for poor black children. Hamblin, supra note 137, at 11.

157. For instance, scientists have determined that pesticides might be implicated in the rise in children's neurodevelopmental disorders. Cabrera, supra note 143, at 606.
advance that goal. Various domestic laws and international agreements demand a right to a healthy environment, including one free of neurotoxins, but the results are imperfect. As the technology becomes more accessible and less expensive, litigants using neuroscience can be the catalysts to show just how detrimental exposure to these toxins can be. Neuroscience evidence could bolster a possible public nuisance action for use of pesticides, metals, and other chemicals: whereas in the past the consequences of toxins in the environment might be felt too late and a litigant could only recover if he or she had an increased risk of future harm, neuroscience can show effects in the brain in real time. We must also ensure that these technologies are as accessible as possible to all; that detecting, predicting, and screening for neurotoxicity is not prohibitively costly for those with lower incomes; and that the public receives better education on potentially harmful substances.

III. Challenges to and Weaknesses of Neuroscience’s Use in the Law

While neuroscience technology may lead to a fairer legal system, it is still evolving, and despite how tempting it is to draw conclusions from its seductive, technical medical data and graphs, it

158. In the European Union, for example, if a chemical is deemed potentially neurotoxic from testing, it will be strictly regulated. Those regulations “can [later] be relaxed if subsequent testing shows less harm than initially anticipated.” Cabrera, supra note 143, at 609.


160. See, e.g., United Nations Conference on the Human Environment, G.A. Res. 27/2994, U.N. Doc. A/RES/27/2994 (Dec. 15, 1972) (issuing a statement “[r]eaffirming the responsibility of the international community to take action to preserve and enhance the environment and, in particular, the need for continuous international co-operation to this end”).

161. For example, chlortpyrifos, which can affect fetal brain development, is classified as “very highly toxic” to birds and fish, and “moderately toxic” to mammals, but is still used widely in agriculture, greenhouses, wood products, and golf courses. Hamblin, supra note 137, at 4.
cannot yet be relied on as a mind reader, predictor of future actions, or litmus test. While understanding the circuitry and cognition processes of the brain may lead to increased knowledge of human behavior, neuroscience, like any other scientific field, does not always isolate a specific cause that leads to a specific effect. Indeed, the potential value of neuroscience to improve decision-making accuracy and to advance justice must be reconciled with the potential for exaggeration, hype, and premature application of scientific theses that are not yet repeatedly validated. Moreover, incorporating neuroscience into the law is not without its ethical and policy concerns, raising apprehensions regarding conceptions of free will, mental privacy, and personal liberty. Additionally, traditional dichotomies in tort recovery might become superfluous if litigants can demonstrate suffering from emotional injuries using brain scans.

The use of neuroscience technology in the courtroom and in legal policy presents two main categories of concerns: one regarding the reliability and readiness of this technology for use as evidence, and another encompassing the normative, ethical, and policy concerns we might have about the use of this technology. Sections A and B of this Part examine each of these in turn. Section C explains why, despite these concerns, neuroscience is ultimately a useful tool for policymakers and potential litigants. This Section also proposes a number of strategies for legislatures, courts, and society to take to regulate the use of neuroscience data in civil courtrooms.

A. Evidentiary Concerns

From an evidentiary standpoint, it is vital to consider criticisms and weaknesses of utilizing neuroscientific data in the courtroom. Primarily, neuroscience must wrestle with meeting a
threshold of reliability before courts can accept its use in factfinding. 165 Some of those problems include a lack of a baseline, extrapolating information gleaned in generalized studies to a specific instance, confounding social and environmental factors that might influence the data, and the unknown rate of false positive and false negatives. This Section will address each of these limitations.

A significant problem in using neuroimaging evidence is establishing a plaintiff's baseline brain function. 166 For example, without some evidence of an individual’s condition prior to an incident, it is hard to evaluate whether a particular incident actually caused the individual psychological harm or aggravated it further, or whether the individual was suffering from a pre-existing condition. It is unlikely that a plaintiff will have had previous brain scans to compare to the current scan. 167 Similarly, brain scans taking place long after a particular incident occurred may be of limited diagnostic or forensic use. 168 Although the plaintiff may still be experiencing injury or harm, a number of other causes between the injury and the scan could have contributed to the neuroimaging results.

Establishing a baseline goes to the issue of causation; that is, whether a plaintiff's harm is really due to the event in question. 169

165. Scientific consensus on how to interpret the relationship between the observable BOLD response in fMRIs and conclusions regarding mental states of subjects is “still evolving.” Eggen & Laury, supra note 41, at 302.

166. In some cases, as reliance on neuroscience increases and gains popularity, the lack of an individual baseline scan to compare to a post-injury scan may cease to be as problematic. In fact, the NFL and NCAA now give baseline neurological exams to prospective players before they ever play in a game. Carl Zimmer, The Brain: What Happens to a Linebacker’s Neurons?, DISCOVER MAG. (Aug. 18, 2010), http://discovermagazine.com/2010/jul-aug/18-brain-what-happens-to-a-linebackers-neurons [https://perma.cc/6XS2-HD57].

167. Perhaps institutionalizing periodic brain scans for professional and amateur athletes would be one way to implement such an idea. Alternatively, other measures can help paint a “before” picture, such as circumstantial evidence including school, employment, and medical records.


169. In tort law, in order to successfully recover monetary damages, a plaintiff must not only demonstrate an injury, but also that the defendant's action
Although neuroscience has made great strides, we still possess limited understanding of the physical link between brain activity and behavior and must continue to emphasize the distinction between correlation and causation when using neuroimaging data in court. Though fMRIs can accurately measure changes in oxygenated blood flow, interpreting those changes as reliable indicators of particular types of thought, or as reliable indicators of what a region of the brain is actually doing, requires a series of inferential steps that involve statistical analysis, interpretation, and comparison to other information. In fact, studies from psychology, psychiatry, and public health have shown that a traumatic event may be a cause of a mental disorder but may not be the proximate or sole cause.\(^{170}\)

As insightful as brain scanning techniques are, it is crucial to remember that they are proxy measures of brain activity. There is substantial “human judgment” between data acquisition and the creation of the “eye-catching fMRI images that we have become accustomed to seeing.”\(^{171}\) Experts’ analyses can help interpret neuroscience data, but they can also lead to distortions of it. In the courtroom, judges must consider the credibility of the neuroscience evidence, the ways in which it might be interpreted or manipulated, and its potential impact on jurors.

Another criticism of reliance on this technology is that there is a significant difference between how a brain functions in laboratory experiments and how a brain experiences the real world in the midst of an incident.\(^{172}\) It may be difficult, if not impossible, to approximate the real-world context during a brain scan. Additionally, most of what we know about brain function comes from studies that average results from groups of individuals; as a result, it is challenging to predict the exact nature of brain dysfunction in individual subjects. Brains and their responses vary not just across individuals but also within particular individuals over time due to external contextual factors like

\(^{170}\) Oliver R. Goodenough & Micaela Tucker, Law and Cognitive Neuroscience, 6 ANN. REV. L. & SOC. SCI. 61, 66 (2010) (arguing that the brain is a composite of influences incorporating numerous social, cultural, and personal experiences; even medication can alter fMRI signals).


\(^{172}\) Many neuroscience studies are done on compliant test subjects, often college students. Cost, the availability of volunteers, and ethical and practical hurdles make it difficult to generate more realistic or diverse studies. Jay Aronson, The Law’s Use of Brain Evidence, 6 ANN. REV. L. SOC. SCI. 93, 100 (2010).
mood, medication, or sleep deprivation. It is imperative to consider false positives and false negatives: a person could subjectively experience pain because of a low pain threshold even though it may not manifest on a scan, or she could feel no pain while a scan indicates that she should be feeling it.

Keeping these limitations in mind, we must weigh the probative value of neuroscience evidence against potential prejudicial impact on judges and juries. In other words, fact finders may tend to trust brain scans simply because they are impressed by the images and believe them to be scientifically objective. According to Federal Rule of Evidence 403, the court may “exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.” Colorful neuroimaging in a courtroom may seem more reliable to a jury than is justified. On the other hand, this is nothing new: all scientific evidence, particularly when it comes from more novel technologies, can lead to confusion and is rarely, if ever, 100% dispositive.

Neuroscience technologies are still developing and require further corroboration before they are ready for widespread use in the

173. Greely, supra note 18, at 182 ("Brains are complicated and individual . . . It is unlikely that everyone's brain will react the same way to exactly the same stimulus.").


175. While it is difficult to estimate how influential neuroscience evidence can be in the courtroom, Nita Farahany found that in cases where defendants used neuroscientific evidence, they received a slightly more favorable outcome. Greg Miller, The Brain Gets Its Day in Court, ATLANTIC (Mar. 1, 2016), https://www.theatlantic.com/science/archive/2016/03/neurolaw-brain-scans-court/471615/ [https://perma.cc/3GYZ-LZSK].

176. FED. R. EVID. 403. Rule 403 has been invoked to exclude evidence from polygraph tests. Greg Miller, Brain Scans of Pain Raise Questions for the Law, 323 SCI. 195 (2009).

177. Hank Greely points to the forensic use of DNA, unreliable eyewitnesses, and even the introduction of PowerPoint slides as examples of controversial or imperfect pieces of evidence in the past. Seminowicz et al., supra note 95, at 230–31. Moreover, there is evidence that neuroscience evidence may not be as excessively persuasive as some believe it to be. Nicholas J. Schweitzer et al., Neuroimages As Evidence in a Mens Rea Defense: No Impact, 17 PSYCHOL. PUB. POLY & L. 357, 366 (2011) (finding no evidence that neuroimaging unduly influences juries over verbal neuroscience-based evidence).
courthouse. Ensuring scientific reliability, commissioning larger and more diverse sample pools, developing theories to help explain and test correlations, and increasing scientific literacy among decisionmakers will contribute to increasing neuroscience's legal value and to decreasing skepticism. Still, it is important to note how advanced brain scan technology has become—for example, scientists are now able to evaluate specific structures, chemical levels, and individual receptors in the brain.\textsuperscript{178} Perhaps the best way to approach the power of these technologies is to use neuroscience as only one factor among others—as a piece of evidence added to the whole puzzle that the triers of fact consider in reaching a decision. Neuroscientific data alone may not always be conclusive but could be corroborated with litigants' self-reports and behavior.\textsuperscript{179}

While the limitations outlined in this Section may support evidentiary challenges or grueling cross-examination, they do not justify an absolute legal barrier to the use of brain scan technology in the legal arena, especially considering the probative value that the data can provide. Although neuroscientific imaging methods may never be infallible or definitive enough to supersede other forms of evidence, there are many suffering people to whom these methods can finally give credence and validation.

B. Normative, Ethical, and Policy Challenges to Neuroscience in the Courtroom

Science may inform the law but it does not dictate it. As a society, we may want to maintain distinctions between physiologically similar harms for normative reasons. Courts are skeptical of recognizing invisible injuries for several reasons other than the evidentiary difficulties discussed in Section A. They may fear that eliminating the distinction between awarding economic damages for physical but not invisible injuries is the first step in a slippery slope.\textsuperscript{180}

\textsuperscript{178} Greely, supra note 18, at 181.

\textsuperscript{179} One could envision, for example, an expert using generalized neuroscience statistics as a guide against an individual's data and say something like, "when we see this pattern of brain activation, in similar circumstances, 90 percent of people we believe to be honest report that they're in pain" rather than conclusively decide that the data says something certain about the individual. See Greely, supra note 18, at 182–83.

\textsuperscript{180} See, e.g., Metro-North Commuter Railroad Co. v. Buckley, 521 U.S. 424, 442 (1997) (expressing concerns about recognizing "unlimited and unpredictable" liability and rejecting the plaintiff's claims to damages and monitoring services on
A number of scholars defend the relevance of the distinction between physical and emotional harm, especially in tort law, by asserting that a duty to maintain one's own emotional well-being can benefit both tort plaintiffs and defendants because it incorporates normative ideals about identity, consent, autonomy, social justice, and social welfare. After all, not everything that we "dislike or resent, and wish to avoid, is harmful to us," including unpleasant but not unequivocally debilitating mental states such as disappointment, hurt feelings, broken hearts, and shame. Not all stress is bad and negative experiences or emotions can be helpful and constructive; these types of injuries are inherent in the experience of being human and may not be worthy of compensation in the courtroom.

Additionally, the use of neuroscience technology in the courtroom or its incorporation into legislation could backfire and hurt some litigants. It is imperative to consider whether we, as a society, want to develop new standards of objectively measuring injury or harm. While using brain imaging technology might allow compensation for some litigants who would otherwise be unable to demonstrate an objective measure of their pain and suffering, a reliance on these technologies might exclude other litigants who are unable to show this harm on a device. As neuroscience evidence becomes a norm in the courtroom, litigants unable to proffer this evidence—whether due to cost or other involuntary means—may be disadvantaged or their lawyers charged with ineffective assistance of counsel.

the grounds that he could only recover them if and when he manifested symptoms of a disease).

181. Erica Goldberg, Emotional Duties, 47 CONN. L. REV. 809, 811 (2015) (maintaining that the distinction between physical and emotional harm "should be based on a duty that we all have to reasonably regulate our own emotional well-being").


183. For example, increased reliance and insistence on brain scanning techniques in litigation might even prolong PTSD. L. H. Field, Post-traumatic stress disorder: A reappraisal, 92 J. ROYAL SOC'Y MED. 35, 35 (1999) ("[O]ngoing litigation acts as an artificial reinforcing factor for unpleasant memories and their accompanying affect.").

184. For example, some litigants might be unable to show harm due to false positives or false negatives. See supra Part III.A.

If some type of neuroimaging evidence is admissible and becomes expected, can a litigant be penalized for not proffering such evidence? In one case, the United States Court of Appeals for the Eighth Circuit suggested that a plaintiff's expert should have ordered a PET or SPECT scan of the plaintiff's brain to support a PTSD claim.\textsuperscript{186} Because the expert failed to do so, the court rejected the argument that the plaintiff suffered physical injury to her brain. It is not farfetched to imagine a future where defense counsel routinely requests fMRI tests or requires that a plaintiff be made available for an fMRI, as is currently done in DNA or other forensic tests, to substantiate claims of pain and suffering.

Increased use of neuroscience data in the courtroom could also backfire by creating a problematic perception that a claimant has a duty to mitigate his or her own harm. Might victims of tortious wrongdoing be required to avert the aggravation of their own injuries? Expecting invisible injury victims to, say, go to a therapist or take medications or painkillers, and penalizing them with a lower damage award if they do not, could infringe upon cognitive liberty and other societal or cultural notions of autonomy.

Socioeconomic obstacles to using this technology to boost litigants' claims are also concerning. The cost of neuroimaging scans may be prohibitive for some tort litigants.\textsuperscript{187} A preference for scans might prejudice decisionmakers against claimants who cannot afford the technique or whose condition cannot reliably be discerned by a scan.

Neuroscience and its potential to eliminate the physical-mental divide in our legal system can also raise a number of constitutional issues, possibly implicating the First, Fourth, Fifth, Seventh, Eighth, or Fourteenth Amendments. First Amendment doctrine relies on the notion that speech that causes emotional harm should be less susceptible to regulation than conduct that causes physical harm, a norm that neuroscientific insights might challenge. A brain scan could also constitute a search of the person, implicating the Fourth Amendment.\textsuperscript{188} The Fifth Amendment protects individuals

\textsuperscript{186} Lloyd v. American Airlines, 291 F.3d 503, 511 (8th Cir. 2002).

\textsuperscript{187} Costs for fMRI, PET, and other scans vary by region and by insurance plans. According to one source, an fMRI scan costs $539 per hour. Yale Sch. of Med., Usage Charges, MAGNETIC RESONANCE RES. CTR. (July 1, 2018), http://mrrc.yale.edu/users/charges.aspx [https://perma.cc/WYK4-EAG7].

\textsuperscript{188} Rosen, supra note 36 (asking whether police can “get a search warrant for someone’s brain”).
from being forced to incriminate themselves—can a brain scan disrupt that notion? To the extent we have a right to mental privacy and cognitive liberty, who should be entitled to access this information, and is this right protected in the Constitution? The ways use of neuroscience in the courtroom could implicate the right of trial by jury, protected by the Seventh Amendment, and due process and equal protection rights, protected by the Fourteenth Amendment, are numerous and complex, and likely will become relevant as its use in the courtroom becomes more prevalent. Finally, punishing people for their neurobiological thoughts rather than for their actions could violate the Eighth Amendment’s ban on cruel and unusual punishment. These questions are beyond the scope of this Note, but remain important to consider.

C. The Way Forward: Creating Law and Policy Informed by Neuroscience Research

Although neuroscience research and tools will never provide all of the answers about a given case or person, its probative value outweighs the risks mentioned in Sections A and B. Insights from functional neuroimaging evidence are increasingly illuminating, relevant, and reliable, and they can at least supplement more conventional evidence. Neuroscientific data can increase confidence in the law’s conclusions and in some cases challenge our confidence in those conclusions, and it can inform how we define tortious behavior. Neuroscience and law will inexorably continue to intersect as our understanding of the brain becomes more sophisticated and as lawyers become more familiar with neuroscientific evidence’s potential. Therefore, it is important to discuss regulatory, judicial, and doctrinal options for neuroscience in the legal field.

Because careful consideration is required to determine when the use of neuroscience evidence is appropriate, it might be helpful to establish a working group or commission to develop rules of evidence to accommodate the technological developments. There is already an organization focused on using neuroscientific insights to inform legal policy in the criminal law context, but perhaps one focused on civil

189. The idea of holding people accountable for their predispositions rather than their actions “poses a challenge to one of the central principles of Anglo-American jurisprudence: namely, that people are responsible for their behavior, not their proclivities—for what they do, not what they think.” Id.

and tort contexts should be established as well. This group could include neuroscientists, psychologists, economists, and lawyers, among others, and should be socioeconomically inclusive so that all parts of society are allowed equal input. This group could discuss the various ways that neuroscience might be used and especially how it might be misused in the courtroom. Researchers and scientists could present their findings and explain the limitations of their findings (such as making it clear that a study consisted of, say, mostly college-age students and that an averaged data point might not extrapolate to a particular individual); behavioral psychologists and economists could interpret these findings and offer predictions about how they might play out in society; and lawyers could offer their perspective on how a judge or jury might consider and respond to that data. Policymakers could then propose model laws to accommodate these pieces of evidence or perhaps offer guidelines for courts interpreting neuroscience evidence in Daubert proceedings. It might be beneficial to craft a standard for the use of neuroscience evidence in court and for policy determinations similar to the standards used for genetic data or forensic testing and to consult methods employed by groups such as the National Conference of Commissioners on Uniform State Laws, including research, drafting, and oversight committees.

In judicial proceedings, consulting neuroscience and cognitive psychology experts will be crucial. Under federal evidentiary standards, neuroscience data should be viewed with caution, but should still be considered as one of the factors in an overall evaluation of harm. The evidentiary concerns listed in Section A can be mitigated by instituting clear jury instructions on how to analyze neuroscience evidence objectively and meticulously. As long as juries and judges are instructed to consider neuroscience data with a critical eye, neuroscience evidence should be accepted in court as substantiation of invisible injuries and it should be viewed as helpful, though not dispositive. A variety of indicators could be used to evaluate the neuroscience evidence—such as how reliable the data is or who exactly

---

191. One possible avenue would be for the National Conference of State Legislatures (NCSL) to publish a uniform state law regarding how to interpret neuroscience research in a courtroom.

the data comes from—and could thereby inform a court or a jury on how much weight they should give the evidence.

Courts and legislatures will increasingly be called upon to articulate new boundaries of liability in the area of invisible harms, but so will society at large. At the very least, neuroscience will force a reconsideration of what some see as outmoded distinctions in the law. As Betsy Grey asserts, “if there is to be no or lesser recovery for mental distress claims, then this choice should be better explained by policy concerns about ruinous liability and a desire to reserve funds for victims of other harms rather than based on an unexamined physical-mental boundary.”193 Furthermore, taking into account how research has chronicled the damaging neurobiological effects of harms such as PTSD, mTBIs, and neurotoxins, neuroscience might change what we see as reasonably foreseeable for the purposes of determining proximate causation in tort law. If neuroscientists can show a strong correlation between these harms and neurological and psychological injuries, plaintiffs may find the causation prong of a tort claim easier to prove.

Ultimately, we may not want to remedy every harm that one can experience in society. But more objective, measureable information that is rooted in the physiology of invisible injuries could reveal that some actions injure more than others. We might want to protect against these actions by redrawing the lines in our tort system. Neuroscience might be able to tell us what traditionally unprotected tortious harms can cause long-term damage or have other harmful effects. For example, a physical disability or injury is generally seen as more debilitating than a bully’s harsh words, but what if a reliable brain scan showed that those words affected a particular individual so greatly that he could not attend school, sleep, or socialize?194 While we might not want to bring hurt feelings or broken hearts into litigation, neuroscience shows that words, in some instances, can hurt in a quantifiable way.195 Once neuroscience technology is able to verify injuries like this—and the technology is close if not there already—tort law should depend less on the arbitrary physical-mental divide and

194. “If the plaintiff is to recover every time that her feelings are hurt, we should all be in court twice a week... But this is a poor reason for denying recovery for any genuine, serious mental injury. It is the business of the law to remedy wrongs that deserve it, even at the expense of a ‘flood of litigation.’”). William L. Prosser, Intentional Infliction of Mental Suffering: A New Tort, 37 MICH. L. REV. 874, 877 (1939).
195. Eisenberger, supra note 81, at 42.
focus more on an individual’s particular experience of harm, its severity for the particular litigant, and the extent to which the harm can be attributed to the offender’s conduct.

It may be difficult to assess damages for some of these torts, and financial compensation alone cannot necessarily restore an invisibly injured plaintiff. However, financial compensation can still serve the important purpose of reifying social norms against the harm and affirming a plaintiff’s bodily integrity. Indeed, the insights from this technology can help, protect, and dignify people who experience injury and impairment, especially for those who previously have never been given such a platform.

CONCLUSION

Just as neuroscience technology can be used to rehabilitate or sentence defendants more fairly in criminal law, it can also yield a more just allocation of resources in civil and human rights law. Not only will it provide more objective evidence for invisible injuries, but it can also provide due process to those who traditionally have not had access to courts, particularly for those in more vulnerable populations. Despite some critical limitations, the aggregate of insights produced by neuroimaging is impressive and is only improving in reliability. More studies, with more diverse populations and greater attention to possible countermeasures, as well as some testable predictive theories about expected activation patterns, would greatly increase confidence

196. Nita A. Farahany, The Costs Of Changing Our Minds 12 (June 8, 2014) (unpublished manuscript) (on file with the Columbia Human Rights Law Review) (noting that tort compensation for the emotional suffering resulting from rape, for example, “recognizes the enduring impact on victims, and expresses social condemnation of the act and its consequences”).

197. Tommy Jarrett, who suffered from PTSD and lost wages after witnessing a young girl die in a car accident, is one such litigant whose day in court helped restore his life and dignity. Invisibilia: Emotions, NAT’L PUB. RADIO (June 1, 2017), https://www.npr.org/templates/transcript/transcript.php?storyId=530928414 [https://perma.cc/ZA4N-E9K4] (“[E]motional distress is the same thing as physical damage. It can wreak havoc on somebody’s life, and it can destroy them.”).

198. See generally Gertner, supra note 10 (examining how neuroscience can help develop a more informed sentencing approach and reduce the rate of incarceration).

199. If neuroscience “does eventually provide significant insights into the mind, it may well be necessary to revamp our thinking on the Anglo-American system of criminal justice and perhaps our approach to the law entirely.” Elizabeth Bennett, Neuroscience and Criminal Law: Have We Been Getting It Wrong for Centuries and Where Do We Go from Here?, 85 FORDHAM L. REV. 437, 451 (2016).
in neuroscience data. Additionally, suggestions like the ones made above—for a working group to propose model laws or ways to interpret neuroscience evidence in the courtroom—are concrete ways to incorporate insights from neuroscience into the law in a sensible and scrupulous manner.

Neuroscience can offer a better understanding of human behavior and the potential for improved policymaking, increased accuracy, and decreased oversights in advancing justice. Its insights challenge our traditional tort doctrine and policy, forcing us to clarify our reasons for allowing or barring compensation in various contexts, and by extension, which human rights our society most values. Advances in neuroscience help eliminate preconceived assumptions about invisible injuries and suggest that failure of proof is no longer a sufficient excuse to cling to old, outdated doctrines. If tort law is about rectifying a harm inflicted upon someone in society, and as a society we value not only physical but also emotional and mental well-being, then this distinction undervalues the amount of pain and suffering that some people feel and it should be revisited.

---

200. To echo Judge Green, "For Daniel to be left without a remedy under all the undisputed facts in this case is antithetical to the general policy of tort liability in Anglo-American jurisprudence: those who are legitimately injured due to the act or omissions of others should have a remedy in our courts." Ware v. ANW Special Educ. Coop., 180 P.3d 610, 622 (Kan. Ct. App. 2008) (Green, J., dissenting).
Towards new human rights in the age of neuroscience and neurotechnology

Marcello Ienca and Roberto Andorno

Abstract

Rapid advancements in human neuroscience and neurotechnology open unprecedented possibilities for accessing, collecting, sharing and manipulating information from the human brain. Such applications raise important challenges to human rights principles that need to be addressed to prevent unintended consequences. This paper assesses the implications of emerging neurotechnology applications in the context of the human rights framework and suggests that existing human rights may not be sufficient to respond to these emerging issues. After analysing the relationship between neuroscience and human rights, we identify four new rights that may become of great relevance in the coming decades: the right to cognitive liberty, the right to mental privacy, the right to mental integrity, and the right to psychological continuity.

Introduction

The quotation in the epigraph is from the play *Comus*, written by John Milton in 1634. The piece, an exhortation to virtue, follows the story of a young noblewoman who has been abducted by a sorcerer called Comus. He has bounded her to an enchanted chair and tried to seduce her with arguments about the charm of bodily pleasure. Despite all his rhetorical assaults, the woman repeatedly refuses his advances and claims that, no matter what he does or says, she will continue to assert her freedom of mind, which is beyond his physical power. In the end, she is rescued by her brothers, who chase off Comus.

The quoted sentence conveys the idea that the mind is a kind of last refuge of personal freedom and self-determination. While the body can easily be subject to domination and control by others, our mind, along with our thoughts, beliefs and convictions, are to a large extent beyond external constraint. Yet, with advances in neural engineering, brain imaging and pervasive neurotechnology, the mind might no longer be such unassailable fortress. As we will explain in this paper, emerging neurotechnologies have the potential to allow access to at least some components of mental information. While these advances can be greatly beneficial for individuals and society, they can also be misused and create unprecedented threats to the freedom of the mind and to the individuals’ capacity to freely govern their behavior.

In the research context, brain imaging techniques are widely used to understand the functioning of the human brain and detect the neural correlates of mental states and
behavior. Clinical applications of brain imaging as well as other neurotechnologies are significantly improving the well-being of patients suffering from neurological disorders, offering new preventive, diagnostic and therapeutic tools. Outside the clinics, pervasive commercial applications are rapidly providing new possibilities for self-quantification, cognitive enhancement, personalized communication and entertainment for normal users. Furthermore, a number of neurotechnology applications are becoming of major interest in the legal domain, especially tort law, criminal law and law enforcement.

On the other hand, these same technologies, if misused or inadequately implemented, risk creating unparalleled forms of intrusion into people’s private sphere, potentially causing physical or psychological harm, or allowing undue influence on people’s behavior.

This paper makes the case that the possibilities opened up by neurotechnological developments and their application to various aspects of human life will force a reconceptualization of certain human rights, or even the creation of new rights to protect people from potential harm.

In 2013, US President Obama called attention to the potential impact of neuroscience on human rights, emphasizing the need to address questions such as those

“(…) relating to privacy, personal agency, and moral responsibility for one’s actions; questions about stigmatization and discrimination based on neurological measures of intelligence or other traits; and questions about the appropriate use of neuroscience in the criminal-justice system” (Presidential Commission for the Study of Bioethical Issues, 2014).

This article begins by exploring the current possibilities and challenges of neurotechnology, and considers what neurotechnological trends will drive this ethical and legal reconceptualization. After carefully analyzing the relationship between neuroscience and human rights, this paper identifies four new rights that may become of relevance in the coming decades: the right to cognitive liberty, the right to mental privacy, the right to mental integrity, and the right to psychological continuity.

The neurotechnology revolution

For a long time, the boundaries of the skull have been generally considered the separation line between the observable and unobservable dimension of the living human being. In fact, although primitive forms of neurosurgery used in ancient societies, including pseudo-scientific procedures such as trepanation, could allow for the observation and even manipulation (e.g. selective removal) of brain tissue, yet the neural and mental processes run in the brain and underlying emotions, reasoning and behavior remained at length unobservable. In contrast, modern advancements in neuroscience and neurotechnology have progressively allowed for the unlocking of the human brain and provided insights into brain processes as well as their link to, respectively, mental states and observable behavior. In 1878 Richard Canton discovered the transmission of electrical signals through an animal’s brain. Forty-six years later, the first human electroencephalography (EEG) was recorded. Since then, a neurotechnological revolution has taken place inside and outside the clinics. In the 1990s, sometimes referred to as the ‘decade of the brain,’ the use of imaging techniques for neurobehavioral studies increased dramatically (Illes 2003). Today, as a wide and rapidly expanding spectrum of neuroimaging technologies has become clinically and commercially
available, the non-invasive recording and display of patterns of brain activity (often associated with the completion of physical or cognitive tasks) has become standard practice. For example, EEG recordings are being widely used to non-invasively measure electrical activity of the brain and detect voltage fluctuations. In addition, derivatives of the EEG technique such as evoked potentials (EPs) and event-related potentials (ERP) allow to average EEG responses to the presentation and processing of stimuli, hence to record brain signals during the performance of specific sensory, cognitive or motor processes. Another technique, functional magnetic resonance imaging (fMRI), allows to measure brain's electrical activity indirectly, i.e. by using hemodynamic responses (cerebral blood flow) as indirect markers. Current fMRI techniques can localize brain activity, graphically display patterns of brain activation, and determine their intensity by color-coding the strength of activation. fMRI techniques are implemented for a variety of purposes including pre-surgery risk assessment, and functional mapping of brain areas to detect abnormalities (e.g. left-right hemispherical asymmetry in language and memory regions) or to observe post-stroke or post-surgery recovery, as well as the effects of pharmacological and behavioral therapies. In addition, a number of neurological conditions including depression and Alzheimer’s disease can now be diagnosed with the use of fMRI (Koch et al. 2012).

The capacity of neuroimaging techniques to map brain functioning has been tested effective also in gaining insights into people’s intentions, views and attitudes. For example, scientists were able to infer from decoded brain activity which actions participants in their trial were intending to perform. The task in question was to decide whether to add or subtract two numbers and to covertly hold their intention for a few seconds. During that delay, it was possible for scientists to determine with 70% accuracy which of two tasks the subjects were covertly intending to perform (Haynes et al. 2007). In another study, participants toured several virtual-reality houses, and then had their brains scanned while touring another selection. By identifying certain patterns of brain activity for each house, scientists were able to determine which houses their subjects had been to before (Smith 2013). Brain scans do not only allow to ‘read’ concrete experiment-related intentions and memories. They appear even able to decode more general preferences. A US study has shown that fMRI scans can be used to successfully infer the political views of the users by identifying functional differences in the brains of respectively Democrats and Republicans (Schreiber et al. 2013). Similarly, men’s frequent preference for sport cars has been correlated with specific functional differences in the men’s vs the women’s brain (Baron-Cohen 2004).

The possibility of non-invasively identifying such mental correlates of brain functional differences is of particular interest for marketing purposes. Over a decade ago, McClure et al. (2004) used fMRI to show functional differences (increased activation in the dorsolateral prefrontal cortex, hippocampus and midbrain) in the brain of people knowingly drinking Coca Cola as opposed to the same people drinking unlabeled Coke. Their results showed that marketing strategies (e.g. the Coca Cola label) can determine different responses in the brain of consumers (McClure et al. 2004). These results have pioneered the establishment of a spin-out branch of neuroscience at the intersection with marketing research called neuromarketing, which has expanded rapidly over the past decade. Today, several multinational companies including Google, Disney, CBS, and Frito-Lay use neuromarketing research services to measure consumer preferences
and impressions on their advertisements or products. In addition, a number of specialized neuromarketing companies including EmSense, Neurosence, MindLab International and Nielsen, routinely apply neuroimaging techniques, mostly fMRI and EEG, but also Steady State Topography (SST) and physiological measurements (e.g. galvanic skin response) to study, analyze and predict consumer behavior. This possibility of mining the mind (or at least informationally rich structural aspects of the mind) can be potentially used not only to infer mental preferences, but also to prime, imprint or trigger those preferences. For example, Neurofocus, an American multinational neuromarketing company recently acquired by Nielsen, tested subliminal techniques with the purposes of eliciting responses (e.g. preferring item A instead of B) that people cannot consciously register (Penenberg 2011). These techniques included embedding stimuli shorter 30 milliseconds, hence under the threshold of conscious perception. In view of these developments, authors have stressed the need to establish ethical and legal standards for neuromarketing practices (Ulman, Cakar, and Yildiz 2015).

Brain imaging techniques were originally developed and are still mostly implemented within the context of clinical medicine and neuroscience research. In recent years, however, a number of neurotechnology applications have made their way onto the market and are now integrated into a number of consumer-grade devices for healthy users with various non-clinical purposes. The umbrella term usually used to encompass all these non-invasive, scalable and potentially ubiquitous of neurotechnologies is “pervasive neurotechnology” (Fernandez, Sriraman, Gurevitz and Ouiller 2015), a notion borrowed from the most widespread notion of pervasive computing. Today, pervasive neurotechnology applications include brain-computer interfaces (BCIs) for device control or real-time neuromonitoring, neurosensor-based vehicle operator systems, cognitive training tools, electrical and magnetic brain stimulation, wearables for mental wellbeing, and virtual reality systems.

Most of these applications use EEG recordings to monitor electrical activity in the brain for a variety of purposes including neuromonitoring (real time evaluation of brain functioning), neurocognitive training (using certain frequency bands to enhance neurocognitive functions), and device control. EEG-based BCIs are being increasingly used as wearable accessories for a number of everyday activities including gaming, entertainment, and smartphone’s remote control. For example, companies Emotiv and Neurosky offer a large assortment of wireless headsets for everyday use that can be connected to compliant smartphones and personal computers (Ienca and Haselager 2016). Brain-control can be used to remotely control several types of devices and engage in several activities including gaming and other forms of entertainment, marketing, self-monitoring and communicating.

The possibility of non-invasive brain control has raised the attention of the mobile communication industry. Several leading companies including Apple and Samsung are incorporating neurogadgets into the accessory assortments of their major products. For instance, iPhone accessories such as the XWave headset already allow to plug directly into compliant iPhones and read brainwaves. Meanwhile, prototypes of next-generation Samsung Galaxy Tabs and other mobile or wearable devices have been tested to be controlled by brain activity via EEG-based BCI (Powell, Munetomo, Schlueter, and Mizukoshi 2013). In the light of these trends, Yuan and colleagues predicted that neurodevices will gradually replace the keyboard, the touch screen, the mouse and the voice command device as humans’ preferred ways to interact with computers (Yuan, Hsieh, and Chang 2010).
Not only neuroimaging devices and BCIs fit into the category of pervasive neurotechnology. Several electrical brain stimulators fit into this category too. Unlike neuroimaging tools, neurostimulators are not primarily used for recording or decoding brain activity but rather for stimulating or modulating brain activity electrically. Portable, easy-to-use, consumer based transcranial direct current stimulation (tDCS) devices are the most widespread form of consumer-grade neurostimulator. They are used in a number of low-cost direct-to-consumer applications aimed at optimizing brain performance on a variety of cognitive tasks, depending on the brain region being stimulated. Recently, transcranial magnetic stimulation (TMS) - a magnetic method used to briefly stimulate small regions of the brain for both diagnostic and therapeutic purposes, has also evolved into portable devices, which resulted effective in the treatment of migraine (Lefaucheur et al. 2014). Finally, an invasive surgical technique called deep brain stimulation (DBS) involving the implantation of a neurostimulator in the ventrointermediate nucleus of the thalamus has obtained FDA approval and is now increasingly used as a treatment for essential tremor, Parkinson’s disease, dystonia and obsessive–compulsive disorder.

In sum, if in the past decades neurotechnology has unlocked the human brain and made it readable under scientific lenses, the upcoming decades will see neurotechnology becoming pervasive and embedded in numerous aspects of our lives and increasingly effective in modulating the neural correlates of our psychology and behavior. While welcoming continuing progress in neurotechnology development, in this paper we argue that the ethical and legal implications of the neurotechnology revolution should be considered early and in a proactive manner. More in detail, we argue that the legal system has to be adequately prepared to deal with the new challenges that might emerge out of emerging neurotechnology, in particular in the context of human rights. As neurotechnology advances, it is critical to assess whether our current human rights framework is conceptually and normatively well-equipped to face the novel challenges arising at the brain-computer-society entanglement, hence to provide simultaneously guidance to researchers and developers while providing protection to individuals and groups.

**Brain technology and the law**

Neuroscience and the law intersect on many levels and on various different issues. This is not surprising. While neuroscience studies the brain processes that underlie human behavior, legal systems are quintessentially concerned with the regulation of human behavior. It is therefore reasonable to claim that both disciplines are destined to be “natural partners” (Goodenough and Tucker 2010). The underlying idea of the new field called ‘neurolaw’ is precisely that better knowledge of the brain will lead to better-designed laws and fairer legal procedures. Examples of potentially legally relevant applications of neurotechnology are numerous. Brain imaging techniques, for instance, might possibly contribute to more evidence-based decisions in criminal justice, from investigation and the assessment of criminal responsibility, to punishment, rehabilitation of offenders, and the evaluation of their risk of recidivism. The tools offered by neuroscience could potentially play also a role in civil law procedures, for example, in the assessment of an individual’s capacity to contract, or of the severity of the plaintiff’s pain in compensation claims. New and more reliable lie detection technologies based on our knowledge of the brain functioning might help to assess the reliability of witnesses. Memory erasure of recidivist violent criminals and of victims of especially
traumatic offences (e.g. sexual abuse) is also mentioned as another possibility opened by our new knowledge of the brain (Goodenough and Tucker 2010).

A possibly game-changing use of neurotechnology in the legal field has been illustrated by Aharoni et al. (2013). In this study, researchers followed a group of 96 male prisoners at prison release. Using fMRI, prisoners’ brains were scanned during the performance of computer tasks in which they had to make quick decisions and inhibit impulsive reactions. The researchers followed the ex-convicts for 4 years to see how they behaved. The study results indicate that those individuals showing low activity in a brain region associated with decision-making and action (the Anterior Cingulate Cortex, ACC) are more likely to commit crimes again within 4 years of release (Aharoni et al. 2013). According to the study, the risk of recidivism is more than double in individuals showing low activity in that region of the brain than in individuals with high activity in that region. Their results suggest a “potential neurocognitive biomarker for persistent antisocial behavior”. In other words, brain scans can theoretically help determine whether certain convicted persons are at an increased risk of reoffending if released.

This prospect evokes Philip Dick’s 1956 science fiction story “The Minority Report”, which was adapted into a movie in 2002. The plot is about a special police unit (“Pre-crime Division”) which is able to identify and arrest murderers before they commit their crimes. The system is believed to be flawless until an officer from that same unit is mistakenly accused of a future murder (Dick 2002). This dystopian scenario, which could result from the new knowledge about the brain, raises important ethical and human rights questions. How much evidence is needed to prove that brain scans are likely to flag only the truly high risk offenders? Can neurotechnology-generated data, which have a probabilistic nature, be straightforwardly applied to predict the criminal behavior of a particular individual? Can these preliminary findings, which were based on a very specific cohort, be generalized to other groups? In any case, it is clear that much more work is needed to ensure the reliability of the technique before authorizing its use by courts, certainly not as a substitute for current methods for dangerousness assessment, but maybe as an additional, complementary tool.

Other brain technologies that may be relevant for the legal system are lie detectors, mental decoders, and brain printers. Lie detectors are devices capable to record and measure brain responses associated with the retrieval of memories, with the purpose of ascertaining the truth-values of statements relative to those memories. Traditional lie detectors, like the polygraph, measure some bodily markers such as blood pressure, heart rate, and muscular reactions. Despite their low reliability, they are regularly used by some government agencies to screen their employees. However, they are very rarely accepted as evidence in US courts. The new generations of lie detectors, which are EEG-based and fMRI-based, are regarded as much more reliable than the polygraph, as they detect the lie at its source: the brain. In the United States, at least two companies -No Lie MRI and Cephos Corp - are currently offering fMRI lie-detection services (Greely 2009). A study published in 2005 by a research group linked to Cephos, claimed that fMRI-based lie detection has a reliability of around 90%. The study predicted that the procedure will be further improved and ready to be used in courts in the not too distant future (Kozel et al. 2005). More recent studies have confirmed the higher reliability of fMRI-based lie detectors compared to polygraphy (Langleben et al. 2016). In parallel, mental decoders are capable of decoding mental states and translating them into observable
outputs such as text, verbal signals or graphic images. For example, Herff et al. (2015) and Mirkovic et al. (2015) have independently demonstrated the effectiveness of a decoder capable of reconstructing speech from brain waves (Herff et al. 2015; Mirkovic, Debener, Jaeger, and De Vos 2015). Such devices have a great potential for clinical applicability as they could benefit several classes of neurological patients, especially those suffering from locked-in syndrome and paralysis. Such patients, who might have lost their capacity to produce verbal communication, would be enabled to re-interact with the external world by producing speech solely by brain activity. Outside the clinical setting, such decoders are tested to enhance mobile communication through thought-to-text converters. Not all mental decoders are designed to enhance users’ autonomy. Some devices are currently tested for monitoring brain states with the purpose of guiding the individual’s behavior.

For example, NASA and Jaguar are jointly developing a technology called Mind Sense, which will measure brainwaves to monitor the driver’s concentration in the car (Biondi and Skrypchuk 2017). If brain activity indicates poor concentration, then the steering wheel or pedals could vibrate to raise the driver’s awareness of the danger. This technology can contribute to reduce the number of accidents caused by drivers who are stressed or distracted. However, it also opens theoretically the possibility for third parties to use brain decoders to eavesdropping on people’s states of mind.

Similar implications are raised by brain printers. These are prototypical devices that are currently tested as brain-based authentication methods. For example, researchers at Binghamton University in the state of New York have devised a way to verify a person’s identity based on how their brain responds to certain words. The researchers observed the brain signals of 45 volunteers as they read a list of 75 acronyms, such as FBI and DVD, and recorded the brain’s reaction to each group of letters, focusing on the part of the brain associated with reading and recognizing words. It turns out that participants’ brains reacted differently to each acronym - so that a computer system was able to identify each volunteer with 94% accuracy (Armstrong et al. 2015). This technology, which could in the short term replace passwords and fingerprints as authentication tool for personal accounts, raises novel privacy and security issues.

As neurotechnology advances and opens novel opportunities for monitoring and controlling cognitive function, there is uncertainty on how the law should cope with such advancements. In particular, it remains debatable whether emerging trends in neurotechnology call for a revision or even a replacement of existing legal concepts at various levels including civil law, tort law, business law and legal philosophy. While increasing attention is being devoted in the literature to emerging neurotechnology applications in the context of criminal law or to the increasing use of neuroscience evidence in courts, little focus has been directed to the implications of advancing neuroscience and neurotechnology for human right law. This neglected component of the neurolaw discourse is of particular relevance since the universal nature of the human right framework could provide a solid foundation for this emerging ‘jurisprudence of the mind’.

**Neuroscience and human rights**

**Overview**

While neurotechnology has the potential to impact human rights such as privacy, freedom of thought, the right to mental integrity, the freedom from discrimination, the right to a fair
trial, or the principle against self-incrimination, yet international human rights law does not make any explicit reference to neuroscience. In contrast to other biomedical developments, which have already been the subject of standard-setting efforts at the domestic and international level, neurotechnology still largely remains a terra incognita for human rights law. Nonetheless, the implications raised by neuroscience and neurotechnology for inherent features of human beings, urge a prompt and adaptive response from human rights law.

The adaptive ability that human rights law has shown in responding to the challenges posed by genetic technology may help to anticipate how this branch of law could evolve in the coming years in response to new issues raised by neuroscience. Since the end of the 1990s, the international community has made significant efforts to address a great variety of issues that result from the increasing access to human genetic data. In 1997, the Universal Declaration on the Human Genome and Human Rights (UDHGR) was adopted to prevent that genetic information is collected and used in ways that are incompatible with respect for human rights, and to protect the human genome from improper manipulations that may harm future generations. The principles contained in this instrument were further developed in 2003 by the International Declaration on Human Genetic Data (IDHGD), which sets out more specific rules for the collection of human biological samples and genetic data. It is interesting to note that from the interaction between genetics and human rights resulted entirely new rights, such as the ‘right not to know one’s genetic information,’ which is formally recognized by the UDHGR (Art. 5(c)) and the IDHGD (Art. 10), as well as by other international and national regulations. In addition to the recognition of new rights, ‘old’ rights -such as the right to privacy and the right against discrimination- were specifically adapted to the novel challenges posed by genetics. This close connection between life sciences and human rights was further strengthened by the 2005 Universal Declaration on Bioethics and Human Rights, which comprehensively addresses the linkage between both fields (Andorno 2013). This latter document sets out principles that are applicable not only to genetics but to other biomedical and life sciences issues.

In this paper we claim that, similarly to the historical trajectory of the ‘genetic revolution,’ the ongoing ‘neuro-revolution’ will reshape some of our ethical and legal notions. In particular, we argue that the growing sensitivity and availability of neurodevices will require in the coming years the emergence of new rights or at least the further development of traditional rights to specifically address the challenges posed by neuroscience and neurotechnology. This argument is in accordance with the observation of how human rights have historically emerged and developed in modern societies. Human rights, in fact, have always arisen as specific responses to recurrent threats to fundamental human interests (Nickel 1987), to human dignity (Habermas 2010), or to what is required by a “minimally good life” (Fagan 2005). As we attempt to show in this paper, the individual quest to exert control over one’s own neuro-cognitive dimension as well as the emergence of potential threats to basic human goods or interests posed by the misuse or inadequate application of neurotechnological devices may require a reconceptualization of some traditional human rights or even the creation of new neuro-specific rights.

It goes beyond the scope of this article to discuss the different theories about the foundations of human rights, or to take a position in this regard. For the purposes of our investigation we chose to adopt a broad practical conception of human rights, like the one proposed by Beitz (2011, p. 109), who argues that they are “requirements
whose object is to protect urgent individual interests against predictable dangers ('standard threats') to which they are vulnerable under typical circumstances of life in a modern world order composed of states’ (Beitz 2011). In general terms, it can be said that the scope of human rights is to guarantee both the necessary negative and positive prerequisites for leading a minimally good life (Fagan 2015).

A common objection against the recognition of new rights is that it leads to the so-called ‘rights inflation’, which is the objectionable tendency to label everything that is morally desirable as a ‘human right’. The unjustified proliferation of new rights is indeed problematic because it spreads skepticism about all human rights, as if they were merely wishful thinking or purely rhetorical claims. Rights inflation is to be avoided because it dilutes the core idea of human rights and distracts from the central goal of human rights instruments, which is to protect a set of truly fundamental human interests, and not everything that would be desirable or advantageous in an ideal world.

A frequently accepted way to avoid rights inflation is to impose justificatory tests for specific human rights. For example, according to Nickel (2014), it could be required that a proposed human right should not only deal with some very important good but also respond to a common and serious threat to that good, impose burdens on the addressees that are justifiable and no larger than necessary, and be feasible in most of the world’s countries (Nickel 2014). The international law scholar Philip Alston (1984) has suggested a list of criteria that a given claim must satisfy in order to qualify as a ‘human right’ in terms of international law. In his view, the proposed new human right must “reflect a fundamentally important social value”; “be consistent, but not merely repetitive, of the existing body of international human rights law”; “be capable of achieving a very high degree of international consensus”, and “be sufficiently precise as to give rise to identifiable rights and obligations”.

For the reasons we give below, we think that the new rights advocated in this paper – the right to cognitive liberty, the right to mental privacy, the right to mental integrity, and the right to psychological continuity – fulfill these requirements and therefore do not raise the risk of rights inflation.

This proposal of neuro-specific human rights is consistent with Glen Boire’s advocacy of a “jurisprudence of the mind” that “takes account of the latest understandings of the brain” and “which situates these within our country’s tradition of embracing individual, self-determination and limited government” (Boire 2003, p. 10). As brain technology is rapidly reshaping the infosphere and the digital infrastructures in our societies, there is an urgent need to proactively assess whether our current ethical and legal frameworks are ready to face this emerging scenario.

At this stage it is also worth noting that many of the issues discussed in this paper are not unique to cutting-edge neurotechnology but have precedents in more traditional interventions. For example, breaches for mental privacy emerged before the invention of neuroimaging and neuromonitoring technologies through more rudimental techniques such as interrogation and polygraph-based lie detection. These interventions, however, do not target neural processing directly but only via proxy-processes such as speech, behavior, and physiological indices (e.g. pulse and skin conductivity). In addition, the degree of accuracy and resolution of such techniques is remarkably low (Iacono 2008), hence often insufficient to support epistemologically justified inferences about mental information. Similarly, threats to mental integrity and psychological
continuity were posed by non-computational interventions such as psychoactive drugs and hypnotic inductions way before the invention of neurostimulation and brain-machine interfacing. However, these techniques are often characterized by limited efficacy and reliability in purposively manipulating mental activity as well as low degrees of selectivity in targeting neural processes. Based on these considerations, we argue that advanced neurotechnology enables a degree of access into and manipulation of neural processes significantly higher than other techniques. Therefore, while we consider the ethical and legal analysis presented in this paper applicable to the entire spectrum of both computational and non-computational brain interventions, we argue that the degree of perturbation of advanced neurotechnology on the current ethical-legal framework is quantitatively higher than non-computational techniques. For this reason we situate neurotechnology as the focus of our proposed normative upgrade.

Cognitive liberty

A first, essential step towards the creation of a neuro-oriented human rights framework has been the recent debate over the notion of cognitive liberty. According to Bublitz (2013), this complex notion, often also referred to as mental self-determination, comprises two fundamental and intimately related principles: (a) the right of individuals to use emerging neurotechnologies; (b) the protection of individuals from the coercive and unconsented use of such technologies. As he concisely put it, cognitive liberty is the principle that guarantees “the right to alter one’s mental states with the help of neurotools as well as to refuse to do so” (Bublitz 2013, p. 234).

Proponents of cognitive liberty suggest considering it a “fundamental human right” as well as “a central legal principle guiding the regulation of neurotechnologies” (Ibid.). The reason of its fundamental function stems from the fact that “the right and freedom to control one’s own consciousness and electrochemical thought processes is the necessary substrate for just about every other freedom” (Sententia 2004). In fact, as Bublitz argued, “it is hard to conceive any conception of a legal subject in which the mind and mental capacities (e.g. acting from reasons, deliberation) are not among its necessary constitutive conditions” (2013, p. 242). Cognitive liberty, therefore, is necessary to all other liberties, because it is their neurocognitive substrate. As such, cognitive liberty resembles the notion of ‘freedom of thought’ which is usually considered the essential justification of other freedoms such as freedom of choice, freedom of speech, freedom of press, and freedom of religion. Not surprisingly, Sententia (2004) presented cognitive liberty as a conceptual update of freedom of thought that “takes into account the power we now have, and increasingly will have to monitor and manipulate cognitive function”. Some legal scholars such as Boire and Sententia have interpreted the right to cognitive liberty with special focus on the protection of individual freedom and self-determination from the State. For example, Sententia has claimed that “the State cannot, consistent with the First Amendment of the Constitution, forcibly manipulate the mental states, and implicitly the brain states of individual citizens”.

Given its conceptual complexity, cognitive liberty is multi-dimensional. Bublitz recognizes at least three “interrelated but not identical dimensions” (Bublitz 2013, p. 251). These are: (i) the liberty to change one’s mind or to choose whether and by which means to change one’s mind; (ii) the protection of interventions into other minds to
protect mental integrity, and (iii) the ethical and legal obligation to promoting cognitive liberty. These three dimensions configure cognitive liberty as a complex right which involves the prerequisites of both negative and positive liberties in Berlin’s sense (Berlin 1959): the negative liberty of making choices about one’s own cognitive domain in absence of governmental or non-governmental obstacles, barriers or prohibitions; the negative liberty of exercising one’s own right to mental integrity in absence of constrains or violations from corporations, criminal agents or the government; and finally, the positive liberty of having the possibility of acting in such a way as to take control of one’s mental life.

Being the neurocognitive substrate of all other liberties, cognitive liberty cannot be reduced to existing rights, hence is immune to the risk of rights inflation. In addition, since cognitive life, although in various forms and degrees, is inherent in all human beings, cognitive liberty is consistent with a definition of human rights as inalienable fundamentals rights “to which a person is inherently entitled simply because she or he is a human being” (Sepuldeva, Van Banning, and van Genugten 2004), regardless of their nation, location, language, religion, ethnic origin or any other status. Consequently, its integration into the human right framework would enable the protection of constitutive features of human beings that are not being entirely protected by existing rights.

For the purposes of our analysis, in this article we will focus exclusively on the negative formulation of the right to cognitive liberty, namely as the right to refuse coercive uses of neurotechnology. In addition, while we welcome the introduction of the right to cognitive liberty, we argue that this notion is not sufficient alone to cover the entire spectrum of ethical and legal implications associated with neurotechnology. Rather, the establishment of cognitive liberty as a human right should be coordinated with a simultaneous reconceptualization of existing rights or even the creation of other new neuro-specific rights. These are the right to mental privacy, the right to mental integrity and the right to psychological continuity.

The right to mental privacy

The right to privacy

Today’s infosphere is more intrusive than at any other time in history. Websites regularly use cookies to record store visitors’ information such as browsing activities, preferences, personal data, visited pages, passwords, credit card numbers, etc. Big and small corporations engage in data-mining activities that capture massive amounts of data about users. Much of this information is about daily activities: what was purchased, when, where and how much was paid. E-mail accounts are stuffed with advertisements and unsolicited offers. Phone numbers and personal addresses are captured in databases and sold to corporations and government agencies. In addition, video surveillance, facial recognition technology, spyware are opening up people’s daily activities for public consumption. As Moore (2010) puts it, today “informational privacy is everywhere under siege”.

The widespread availability of neurotechnology applications will provide multiple opportunities for individuals to access and exert control over their brain-activity, hence resulting in a number of potentially beneficial activities such as self-monitoring, neuro-enhancement, and brain-controlled computer use. However, these same tools will disseminate an unprecedented volume and variety of brain information outside the clinical
domain and potentially increase the availability of such information to third parties. As pervasive applications of neurotechnology are introducing brain data into the infosphere, they are thereby exposing them to the same degree of intrusiveness and vulnerability to which is exposed any other bit of information circulating in the digital ecosystem. At present, no specific legal or technical safeguard protects brain data from being subject to the same data-mining and privacy intruding measures as other types of information. In the words of Nita Farahany, “there are no legal protections from having your mind involuntarily read”. The reason for that stems from the fact, as Charo (2005) observes, that “technology innovates faster than the regulatory system can adapt”.

A large number of ethical, legal, and social questions arise from these neurotechnological possibilities. These include: For what purposes and under what conditions can brain information be collected and used? What components of brain information shall be legitimately disclosed and made accessible to others? Who shall be entitled to access those data (employers, insurance companies, the State)? What should be the limits to consent in this area?

Although a first attempt of response to these questions can be made by appealing to existing legal norms, we claim that specific legal notions and provisions have to be developed. The first notion involved in these debates is that of privacy. International human rights law formally recognises the right to privacy. The Universal Declaration of Human Rights (UDHR) states that “no one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks” (Article 12). Similarly, the 1950 European Convention on Human Rights (ECHR) stipulates that “everyone has the right to respect for his private and family life, his home and correspondence” (Article 8 para 1). It is interesting to note that the right to privacy is one of the few rights that was recognized by international law as a broad, umbrella right before it was included in any state constitution (Diggelmann and Cleis 2014).

At the European level, the right to privacy recognized by the ECHR was developed by the 1995 EU Data Protection Directive (95/46/EC), which specifically aims at protecting individuals with regard to the processing and transfer of personal data. Currently, the EU is planning to adapt the data protection rules to the challenges to privacy posed by the new digital environment. The overall goal of the upcoming Directive and Regulation is to empower individuals with more control over their personal data. Also the EU Charter of Fundamental Rights, adopted in 2000, states the general right to protection of private life in Article 7 and specifies in Article 8 that “everyone has the right to the protection of personal data concerning him or her” (para 1). According to paragraph 2 of the latter provision,"[s]uch data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified”.

The first question that arises in the context of the current privacy protection standards is whether the traditional right to privacy also covers the data contained in and generated by our minds. An answer to this dilemma is not immediately available, not least because there is no consensus in the legal literature on a definition of privacy.
This can be explained by the disparate content of this right, which includes not only the right to control access to personal information, but also to our bodies and to specific private places. In their seminal article, published in 1890, Samuel Warren and Louis Brandeis articulated the right to privacy as “a right to be let alone” (Warren and Brandeis 1890). Their primary concern was the increasing interest of the yellow press in gossiping and revealing personal information about individuals, including pictures of private persons without their consent. This specific instance of privacy was further developed by Alan Westin and other authors into the broader notion of informational privacy, i.e. the control over information about oneself. According to Westin, privacy can be described in terms of our claim to determine for ourselves when, how, and to what extent information about us is communicated to others (Westin 1968). Today, the “right to be let alone” delineated by Warren and Brandeis more than one century ago has clearly become relevant to areas far removed from their original concerns. The various facets of the modern understanding of privacy continue to expand as technological developments continue. Neuroscience is very likely to become in the near future one of the new areas in which the right to privacy is called to play a fundamental and unexpected role.

The emergence of a right to mental privacy

Science fiction can be very helpful to anticipate the challenges that science and technology may pose in the future, as well as the possible responses to them. In a Star Trek novel written in 1990, Captain Kirk has been informed that a dangerous spy has surreptitiously joined one of the groups that are visiting the spaceship Enterprise. Kirk desperately wants to identify the intruder and to know more about him and his plans. By appealing to one of his staff members who has telepath abilities, Kirk wants to read the minds of all the visitors. However, the Captain is reminded by one of his assistants that, according to the law, “the right to mental privacy is an inalienable right of all Federation citizens and shall not be abrogated without due process of law” (Mitchell 1990). Moreover, “to find one guilty individual in either of those groups means there is a large probability of invading the privacy of a number of innocent people” (Ibid., p. 150).

The kind of dilemmas described in this futuristic scenario, which is set in the 23rd-century, may become a reality much earlier than expected. Developments in neuroimaging, like those mentioned above, have raised concerns about the ethics and legality of ‘mind-reading’. It is true that functional brain imaging cannot really ‘read’ thoughts, but can only highlight differences between brain activations during different cognitive tasks, and to infer from such differences certain conclusions about an individual’s thoughts. However, the fact remains that, even if in an indirect manner, these new tools are increasingly able to determine with a high degree of accuracy certain brain data that belong to the private sphere and deserve to be protected from public scrutiny.

In modern societies, privacy and data protection norms cover the use and disclosure of various kinds of personal information. Since the data decoded from an individual’s brain can be regarded as ‘personal information’ – or ‘personally identifiable information’, as it is called in the US –, there is in principle no reason why such data could not be covered by existing privacy and data protection regulations. If one has a “reasonable expectation of privacy” regarding the identifying information derived from one’s blood or saliva samples, surely one has a reasonable expectation of privacy regarding the data decoded from one’s own mind (Shen 2013).
However, the special nature of brain data, which relate very directly to one's inner life and personhood, and the distinct way in which such data are obtained, suggest that specific safeguards will be probably needed in this domain. It should be noted that traditional privacy rules seek to safeguard 'external' information about people.

The particularity of brain data is that the information to be protected is not easily distinguishable from the source itself that produced the data: the individual’s neural processing. This is what we can call the “inception problem”, which complicates the analysis of the issues at stake when traditional approaches to privacy are used. In other terms, the neurotechnological future we are approaching will require us to guarantee protection not only to the information we record and share, but also to the source of that information since they may be inseparable. In order to implement this we would need wider privacy and data protection rights that can be also applied at a higher and chronologically antecedent level: our neural activity.

An additional reason for concern about privacy in this domain is that brain signals allow to distinguish or trace an individual's identity and are potentially linkable to that individual. Some brain records (e.g. EEG-recorded signals) can be used as a unique biometric identifier, similarly to fingerprints or DNA. Back in 2007, Palanippan and colleagues developed a EEG-based biometric framework for automatic identity verification (Palanippan and Mandic 2007). Since then, a huge number of unobtrusive EEG-based biometric systems have been developed for the purposes of individual recognition (Campisi, La Rocca, and Scarano 2012; La Rocca, Campisi, and Scarano 2012), person authentication (Marcel and Del Millan 2007; Palanippan 2008), and person identification (Brigham and Kumar 2010; Mohammadi, Shoushtari, Molae Ardekani, and Shamsoollahi 2006). However, unlike other identifiable information, brainwaves can be potentially recorded without individual's awareness, and therefore in absence of a real ability of the person to consent to the collection and use of that information. With the growing market of portable EEG-based neuroheadsets and in absence of a real possibility for obtaining informed consent for the processing of the records they generate, there is a need for the law to lay down new protective responses to the processing of brain data. The need to protect information generated below the threshold of voluntary control demands for the recognition of a new right that is specifically tailored on the characteristics of brain information and the new possibilities opened by mind-reading technologies.

In the light of the emerging neurotechnologies, it is also necessary to explore the -technical and legal - possibility of applying a filter to the flow of brain information with the purpose of distinguishing the information we consciously want to keep private from the one we want to disclose publicly. In the current information society we are constantly required to draw a distinction between private and public information: for example, when we set up the contact page on our website or when we decide with whom to share our mobile phone number. The basic psychological assumption that underlies this phenomenon is that competent adults have the psychological capacity to consciously filter the information flow and reasonably identify the bits of information that must be kept private. Privacy, in fact, is both a right and an ability. As an ability, it enables individuals or groups to seclude themselves, or information about themselves, and thereby express themselves selectively. This idea has been widely imported into the information technology sphere, where privacy is often described as the ability (or perceived ability) to control submitted personal information -especially when using the
Internet (Dinev and Hart 2004). In order to exercise this ability meaningfully we need a rational medium that is capable to filter the information flow and decide what to disclose. This medium is thought, as well captured by the famous adagio in computer security “the best anti-virus software is the brain”.

Based on these specific challenges, we argue that current privacy and data protection rights are insufficient to cope with the emerging neurotechnological scenarios. Consequently, we suggest the formal recognition of a right to mental privacy, which aims to protect any bit or set of brain information about an individual recorded by a neurodevice and shared across the digital ecosystem. This right would protect brainwaves not only as data but also as data generators or sources of information. In addition, it would cover not only conscious brain data but also data that are not (or are only partly) under voluntary and conscious control. Finally, it guarantees the protection of brain information in absence of an external tool for identifying and filtering that information. In short, the right to brain privacy aims to protect people against illegitimate access to their brain information and to prevent the indiscriminate leakage of brain data across the infosphere.

It is worthy of mention that violations of mental privacy can occur also in absence of direct intrusion into the victim’s neural processing. For example, brain data collected for research purposes are usually stored for analysis on externally located EEG-databases and repositories. Similarly brain-data generated by consumer-grade brain-computer interfaces (BCI) are sent to a connected app and can be stored in the cloud or other data store endpoints. In either case, these data can be accessed also in absence of the person who generated those data and without intervening into the person’s brain signaling.

**Is the right to mental privacy an absolute or a relative right?**

Most human rights, including privacy rights, are relative, in the sense that they can be limited in certain circumstances, provided that some restrictions are necessary and are a proportionate way of achieving a legitimate purpose. In specifically dealing with the right to privacy, the European Convention on Human Rights states that this right admits some restrictions “for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others” (Art. 8, para 2). Only very few rights, such as the freedom of thought, freedom from slavery, torture and inhumane or degrading treatment or punishment are regarded by international human rights law as not subject to any exceptions and, therefore, as absolute rights. In which of both categories should the right to mental privacy be placed? Can nonconsensual intrusions into people’s brain data be justified in certain circumstances or should be unconditionally banned? More concretely, does the right to mental privacy protect individuals from being compelled by courts or the state to brain-based interrogations?

Paul Root Wolpe has suggested that due to fears of government oppression, we should draw a bright line around the use of mind-reading technologies:

“The skull should be designated as a domain of absolute privacy. No one should be able to probe an individual’s mind against their will. We should not permit it with a court order. We should not permit it for military or national security. We should forgo the use of the technology under coercive circumstances even though using it may serve the public good” (Wolpe 2009).
Similarly, it has been argued that “nonconsensual mind reading is not something we should never engage in” (Stanley 2012). The claim is that mind-reading techniques constitute “a fundamental affront to human dignity” (Ibid). Consequently, “we must not let our civilization’s privacy principles degrade so far that attempting to peer inside a person’s own head against their will ever become regarded as acceptable” (Ibid).

Are these calls for an unconditional ban on compulsory mind-reading justified? Or could this procedure be acceptable in certain circumstances (for instance, when faced with a serious crime or a terrorist attack)? As mentioned above, privacy rights are not absolute, but relative. The collection, use and disclosure of private information is permissible when the public interest is at stake. For example, in many jurisdictions, compulsory genetic testing can be undertaken to attempt to identify criminal offenders. Considering the non-invasive and painless nature of brain-scans, there are prima facie good reasons for thinking that their nonconsensual use would be justified, with a court warrant, under special circumstances when there are reasonable grounds to believe that an individual has committed a serious crime or is involved in the planning of a serious crime.

However, this dilemma becomes more intricate when it is seen not in connection to privacy issues, but in the light of the principle of prohibiting coerced self-incrimination. This problem particularly arises when the results of brain scans are regarded not as mere information about individuals (such as buccal or blood-derived DNA, fingerprints, etc.), but as a testimony because in this latter case the self-incrimination clause would enter into play.

The ban on coerced self-incrimination is widely recognized across the democratic world as being an integral component of a fair criminal justice. This privilege is a logic consequence of the presumption of innocence, which places the burden of proof of guilt on the prosecution. In other words, people suspected of a crime do not have any obligation to assist in providing evidence against themselves. The privilege against self-incrimination is very closely related to the right to remain silent and can overlap with it. However, there is a conceptual difference between them: while the former concerns the threat of coercion in order to make an accused yield certain information, the latter concerns the drawing of adverse inferences when an accused fails to testify or to answer questions (Ashworth 2008).

This privilege is enshrined in the International Covenant on Civil and Political Rights, which stipulates that “in the determination of any criminal charge against him, everyone shall be entitled (...) not to be compelled to testify against himself or to confess guilt” (Art. 14(3)(g)). A similar provision can be found in the American Convention on Human Rights and in the Rome Statute of the International Criminal Court. Although the European Convention on Human Rights does not explicitly refer to the privilege against self-incrimination, the European Court of Human Rights (ECHR) has repeatedly asserted that this principle is implied in the general right to a fair trial, which is guaranteed by Article 6 of the Convention. In the US, the Fifth Amendment protects against “coercion [to] prove [a] charge against an accused out of his mouth”. Interpreting this clause, the US Supreme Court introduced in 1966 the distinction between being compelled to provide real or physical evidence (which is allowed) and being forced to give self-incriminating testimony (which is forbidden).

The ECHR draws a more subtle distinction when it differentiates between compelling “real evidence which has an existence independent of the will of the suspect” (ex. documents acquired pursuant to a warrant, breath, blood and urine samples and bodily...
tissues for the purpose of DNA testing) and evidence which is not truly “independent of the will of the suspect”. Answers to questions are the most obvious examples of this second category because they are inconceivable without the will of the subject. However, in the case of Funke v. France, the ECtHR has considered that also being compelled to produce certain documents (in the case, bank statements from accounts in foreign banks, and which might serve to incriminate the individual for tax evasion), would amount to an infringement of the privilege.

Therefore, the lecture of the privilege made by the ECtHR can been understood in the sense that the key issue is not so much whether the evidence is real or oral (i.e. physical as opposed to answers to questions), but whether the evidence requires the active cooperation of the individual or not (Redmayne 2007). In other words, “the privilege only covers assistance from the suspect which could not be substituted by employing direct force” (Trechsel 2005).

If we accept this understanding of the privilege, the question then becomes whether the mere record of thoughts and memories -without any coerced oral testimony or declaration- is evidence that can be legally compelled, or whether this practice necessarily requires the ‘will of the suspect’ and therefore constitutes a breach of the privilege against forced self-incrimination. Unfortunately, it is extremely difficult to give a clear-cut answer to this dilemma. In our opinion, the issue has to be the matter of public discussion in order to find an adequate balance between the private and public interests at stake. The dilemma is particularly arduous because, on the one hand, it could be argued that thoughts and memories are purely internal operations that per se cannot be forced, and consequently the non-incrimination clause would not be applicable to them. However, on the other hand, if mind-reading techniques are allowed in criminal proceedings, there is in the long term the risk to completely water down the privilege against self-incrimination, especially if the techniques become more reliable and efficient than they are at present. People might still be formally protected against self-incriminatory oral statements, but not against the very source of such testimonies: their own thoughts. As Nita Farahany (2012) puts it, self-incrimination may occur silently just as aloud.

The right to mental integrity
Intrusions into people’s brains cannot only result in a violation of their mental privacy, but may also have a direct impact on their neural computation and result in direct harm to them. Ienca and Haselager (2016) have introduced the notion of malicious brain-hacking to refer to neurocriminal activities that influence directly neural computation in the users of neurodevices in a manner that resembles how computers are hacked in computer crime. Focusing on brain-computer interface (BCI), they identify four types of malicious brain-hacking based on the various levels of the BCI cycle where the attack can occur. Three of these types, i.e. when the attack occurs at the level of measurement, decoding and feedback, may involve direct manipulation of a person’s neural computation. Malicious agents may add noise or override the signal sent to the device with the purpose of diminishing or expunging the control of the user over the application, or even hijacking the victim’s voluntary control. For example, a criminal actor could override the signal sent by the users and hijack the BCI-controlled device (e.g. smartphone, electronic wheelchair) without the user’s permission.
In this kind of cases, the users’ mental privacy and the protection of their brain data are not the only rights at risk. Rather, the physical and mental integrity of the victim are at stake too. In fact, the forced intrusion into and alteration of a person’s neural processes pose an unprecedented threat to that person’s mental integrity.

The right to personal physical and mental integrity is protected by the EU’s Charter of fundamental rights (Article 3), stating that “everyone has the right to respect for his or her physical and mental integrity.” Understandably, the Charter emphasizes the importance of this right in the fields of medicine and biology, because of the direct impact that biomedical technologies may have on people’s physical and mental integrity. The provision focuses in particular on four requirements: free and informed consent, the non-commercialization of body elements, and the prohibition of eugenic practices and human reproductive cloning. No explicit reference is made to neurotechnology-related practices. This silence is understandable if we consider that the Charter was adopted in 2000, when the discussion on the ethical and legal implications of neuroscience was at a very early stage. Today however, potential applications of neurotechnology open the prospects of impacting personal integrity in a manner that is comparable to that of genetics and other biomedical practices. For this reason, the normative framework should keep up with neurotechnological advances and extend the protection of people’s integrity to this new area.

We propose to fill this normative gap by calling for a reconceptualization of the right to mental integrity. In fact, while the ECHR and the EU Charter of Fundamental Rights consider mental integrity as a right to mental-health, pendant of the right of physical integrity understood as physical health, a more complex dimension of mental integrity is elicited by neurotechnology. Mental integrity in this broader sense should not only guarantee the right of individuals with mental conditions to access mental health schemes and receive psychiatric treatment or support wherever needed. In addition to that, it should also guarantee the right of all individuals to protect their mental dimension from potential harm.

This reconceptualized right should provide a specific normative protection from potential neurotechnology-enabled interventions involving the unauthorized alteration of a person’s neural computation and potentially resulting in direct harm to the victim. For an action X, to qualify as a threat to mental integrity, it has to: (i) involve the direct access to and manipulation of neural signaling (ii) be unauthorized – i.e. must occur in absence of the informed consent of the signal generator, (iii) result in physical and/or psychological harm. As neurotechnology becomes part of the digital ecosystem and neural computation rapidly enters the infosphere, the mental integrity of individuals will be increasingly endangered if specific protective measures are not implemented.

Threats to mental integrity do not limit to malicious brain-hacking and similar illicit activities. Unauthorized alterations of a person’s neural computation could also emerge out of military applications of BCI technology for warfighter enhancement. Lebedev et al. have described that a neurologically controlled prosthetic could send tactile information back to the brain in nearly real time by using intracortical microstimulation (ICMS), essentially creating a “brain-machine-brain interface” (Lebedev et al. 2011). Such interventions may directly modify neurological activity and can be used to exert some degree of control over ground troop soldiers. For example, the Committee on Opportunities in Neuroscience for Future Army Applications of the National Research Council of the National Academies has investigated the use of portable technologies such as near infrared spectroscopy (NIRS) to detect deficiencies in a warfighter’s
neurological processes and utilizing transcranial magnetic stimulation (TMS) to suppress or enhance individual brain processes (National Research Council 2009). Similarly, mental integrity rights should be included among the rights of war prisoners to prevent the use of invasive brain-washing interventions.

Brain stimulation is an additional domain where the right to mental integrity may play a role. With the growing number of portable neurostimulators available on the market or assembled do-it-yourself devices, the risk that people may misuse these devices with consequent negative impact on their neural functioning should be avoided. For example, while consumer-grade transcranial direct current stimulation (tDCS) are designed to safely function in a certain frequency band, little safeguards prevent users or third persons from manipulating the device’s frequency.

The medical domain is not exempt from the possible application of the right to mental integrity. Invasive neurotechnology interventions such as deep-brain stimulation (DBS) involve the alteration of the patient’s neural processing by electrode-delivered electrical impulses. While this procedure provides therapeutic benefits for otherwise treatment-resistant neurological patients, there is also the potential for neuropsychiatric adverse effects including apathy, compulsive behavior and hallucinations (Mackenzie 2011). In addition, being a surgical procedure, there is a risk of infection, bleeding and rejection of the implanted neurostimulator. Therefore, although in such medical procedure informed consent is always obtained based on minimal medical ethics requirement, still there is a risk that the alteration of neural computation enabled by DBS may cause a disproportionate harm as compared to the therapeutic benefit. This high potential for adverse effects is the reason why, although having proved some effectiveness in the treatment of conditions such as obesity and anorexia nervosa, DBS is still not approved by the Food and Drug Administration (FDA) for the treatment of those conditions. In this context, mental integrity rights stand to prevent from harm, absolutely conceived, but to prevent to a disproportionate relative harm compared to the potential therapeutic benefit.

Finally, the growing field of memory engineering will likely represent a paramount challenge to the right to mental integrity. Several techniques have been developed to engineer (e.g. boost or selectively erase) memories from a person’s mind. For example, Nabavi and colleagues used an optogenetics technique to erase and subsequently restore selected memories by applying a stimulus via optical laser that selectively strengthens or weakens synaptic connections (Nabavi et al. 2014). While they have not reached yet the level of human experimentation, these findings may hold big potential for the treatment of such diseases as Alzheimer’s and post-traumatic stress disorder (PTSD). At the same time, however, the misuse of these techniques by malevolent actors may generate unprecedented opportunities for mental manipulation and brain-washing. For example, criminally motivated actors could selectively erase memories from their victims’ brains to prevent being identified by them later on or simply to cause them harm. On the long term-scenario, they could be used by surveillance and security agencies with the purpose of selectively erasing dangerous, inconvenient from people’s brain as portrayed in the movie Men in Black with the so-called neuralyzer. The potential motives of illicit memory alteration are various, including increasing national security or exerting control over individuals or groups.

Like the right to mental privacy, also the right to mental integrity may not be absolute. For example, it might be argued on utilitarian grounds that controlled and temporary
violations of the right to mental integrity should be allowed as a form of moral enhancement for persistent violent offenders. For example, Persson and Savulescu (2008) have argued that if safe and effective biomedical moral enhancements were developed then they should be compulsory (Persson and Savulescu 2008). Similarly, Ellegaard and Kragh (2015) have argued that it is not only morally permissible, but morally required to force persistent violent offenders to undergo morally enhancing treatments provided the demonstrated effectiveness of such interventions. These possible exceptions to the right to mental integrity would obviously require broad societal discussion to determine whether—and when—such compulsory manipulations of the deepest dimension of the self could be justified for the greater benefit of society.

While taking a position in the long-standing debate over moral enhancement is beyond the scope of this paper, it is important to consider that the postulation of the rights to mental privacy and mental integrity does not ipso facto imply the absolute character of these new rights.

The right to psychological continuity

In addition to mental privacy and mental integrity, also people’s perception of their own identity may be put at risk by inadequate uses of emerging neurotechnology. As we have seen in the first section, neural devices can be used not exclusively for monitoring brain signals but also for stimulating or modulating brain function. For example, transcranial direct current stimulation (tDCS) devices apply constant, low current delivered to the brain area of interest via electrodes on the scalp with the purpose of modulating brain function. Since it causes neuron’s resting membrane potential to depolarize or hyperpolarize, this stimulation causes alterations in brain function that are potentially beneficial for patients. Transcranial magnetic stimulation (TMS) and deep brain stimulation (DBS) open the possibility of intervening into brain function even more substantially. Given the increasing therapeutic effectiveness of tDCS, TMS and DBS, and the rapid advancement of the technology, brain stimulation devices are likely to expand to wider psychiatric groups and, in the case of the first two ones, also to the general population.

However, changes in brain function caused by brain stimulation may also cause unintended alterations in mental states critical to personality, and can thereby affect an individual’s personal identity (Decker and Fleischer 2008). In particular, it has been observed that brain stimulation may have an impact on the psychological continuity of the person, i.e. the crucial requirement of personal identity consisting in experiencing oneself as persisting through time as the same person (Klaming and Haselager 2013). Several cases have been reported in the scientific literature in which DBS has led to behavioral changes such as increased impulsivity and aggressiveness (Frank, Samanta, Moustafa, and Sherman 2007; Sensi et al. 2004) or changes in sexual behavior (Houeto et al. 2002). A study involving patients treated with DBS showed that more than half of them articulated a feeling of strangeness and unfamiliarity with themselves after surgery (“I do not feel like myself anymore”; “I feel like a robot” or “I have not found myself again after the surgery”) (Schüpbach et al. 2006). More recent studies have evidenced personality changes in the direction of increased impulsivity (Lewis et al. 2015; Pham et al. 2015). In parallel, memory engineering technologies may impact a person’s identity by selectively removing, altering, adding or replacing individual memories that are relevant to their self-recognition as persons.
Surely it is an empirical question to determine the frequency and magnitude of these psycho-behavioral changes and it is a question for criminal and tort law to assess the impact of these changes on liability and responsibility. But the question we are interested in here is whether such personality changes induced by neurostimulation or memory manipulating technology could constitute in some circumstances a violation of a basic human right. This might theoretically be the case, for instance, if the patient is legally incompetent (for instance, a child) and the personality change turns out to be psychologically disturbing for him or her. In such circumstances, if the patient’s legal representatives refuse to consent to the removal of the device on the grounds that it has reduced the neurological disorder symptoms, they could be regarded as acting against the individual’s right to psychological continuity.

However, threats to this right are more likely to happen outside clinical settings. For instance, in the context of intelligence and military agencies, it has been reported that over the last decades violations of human rights might have taken place in experiments involving brain electrodes, LSD, hypnosis, the creation of Manchurian candidates, the implantation of false memories and induction of amnesia. Many of these experiments were conducted on unwitting civilians and in the absence of any external review, or representation for the experimental subjects, or any meaningful follow-up (Ross 2007). The new knowledge and technologies in the field of neuroscience clearly offer new and more efficient possibilities for carrying out unconsented personality changes. For example, Pycroft et al. (2016) recently reported the concern that brain implants like DBS are vulnerable to attack by third parties who want to exert malicious control over the users’ brain activity. They called this risk of modification of a person’s brain activity through unauthorized use of neurodevices by third parties ‘brainjacking’ (Pycroft et al. 2016). Negative consequences of brainjacking include (i) information theft, which would result in a violation of the right to mental privacy; (ii) cessation of stimulation, draining implant batteries, inducing tissue damage, and impairment of motor function, which would result in violations of the right to mental integrity. However, some possible consequences of brainjacking such as alteration of impulse control, modification of emotions or affect, induction of pain, and modulation of the reward system could be achieved even in absence of any violation of mental privacy and integrity. In those circumstances of unauthorized modification of the cognitive-emotional-affective dimension a different type of human right violation seems to be at stake: the violation of the right to psychological continuity.

In short, the right to psychological continuity ultimately tends to preserve personal identity and the coherence of the individual’s behavior from unconsented modification by third parties. It protects the continuity across a person’s habitual thoughts, preferences, and choices by protecting the underlying neural functioning. As Paul Tiedemann points out, we understand ourselves as personal unities and as subjects and source of attitudes as long as these attitudes have a minimum level of coherence. This is why a serious lack of coherence makes it impossible to understand oneself (Tiedemann 2016).

The right to psychological continuity can be seen as a special neuro-focused instance of the right to identity. The right to identity was developed by the European Court of Human Rights (ECtHR) from the right to private life included in Article 8 of the European Convention on Human Rights. As we have seen in the first section, Article 8 protects against unwanted intrusion and provides for the respect of an individual’s private space.
However, privacy and personal identity should be distinguished. What the right to psychological continuity aims to prevent is not the unrestricted access to brain information but the induced alteration of neural functioning.

The UDHR also addresses the right to have and develop a personality. Article 22 states: “Everyone is entitled to the realization of the rights needed for one’s dignity and the free development of their personality.” In addition, Article 29 states: “[e]veryone has duties to the community in which alone the free and full development of his personality is possible.” According to Mănuc (2012), personality rights can be defined as those expressing the quintessence of the human person, and are intrinsic to being human. In here analysis, these rights recognize the “spirit” within an individual and have developed from the issues of privacy. It is questionable, however, if current personality rights are well-equipped to address the problem of stimulation-induced alterations in one’s personality.

However, it is questionable whether current personality rights can fully account for the threats posed to psychological continuity. In fact, while this family of rights protects the translation of mental states into action, psychological continuity guarantees protection at an antecedent level: at the level of raw neural functioning. In the risk scenario presented above, misused brain stimulation does not impact the link between mental processes and action, i.e. the expression of mental states, but the mental processes themselves. To provide this more intimate level of protection, there is a need for a new right that preserves the continuity of a person’s mental life from external abusive alteration or disruption.

The right to psychological continuity is closely related to the right to mental integrity, and may factually overlap with it. Both rights stand to protect people from abusive and unconsented alterations of their mental dimension. However, they differ to the extent that the right to psychological continuity also applies to emerging scenarios that do not directly involve neural or mental harm. In contrast, as we have seen in the previous section, the presence of harm is a necessary condition for an action to qualify as an offence to a person’s mental integrity.

To appreciate this difference, it is important to consider that psychological continuity could be threatened not only by misused brain stimulation but also by less invasive, even unperceivable interventions. A good example is unconscious neural advertising via neuromarketing. As we have seen in the first section, neuromarketing companies are testing subliminal techniques such as embedding subliminal stimuli with the purpose of eliciting responses (e.g. preferring item A instead of B) that people cannot consciously register. This has raised criticism among consumer advocate organizations, such as the Center for Digital Democracy, which have warned against neuromarketing’s potentially invasive technology. Jeff Chester, the executive director of the organization, has claimed that “though there has not historically been regulation on adult advertising due to adults having defense mechanisms to discern what is true and untrue”, it should now be regulated “if the advertising is now purposely designed to bypass those rational defenses” (Singer 2010). We argue that a right to psychological continuity can provide the conceptual basis be a viable solution to overcome the problems addressed by Chester.

Potential threats that could be prevented by the right to psychological continuity also include new forms of brain-washing. Holbrook et al. (2016) used transcranial magnetic stimulation (TMS) to neuromodulate the brain regions responsible for social prejudice
and political and religious beliefs. Their results show that by temporarily turning off the posterior medial frontal cortex via TMS it was possible to make participants more positive towards criticisms to their country, than the participants whose brains were unaffected. Using the same technique, they could enhance the participants’ belief in afterlife. While their experiment was designed to mapping the precise neural mechanisms of high-level attitudes and beliefs, their results show that the same technique could be used to trigger a wide spectrum of alterations of a person’s attitudes and beliefs. Malicious agents, for example, could use neuromodulation to exert malevolent forms of mind control. These potentially include religious leaders and coordinators of religiously inspired terrorist groups who want to achieve effective indoctrination and recruitment of youngsters, as well as leaders of authoritarian regimes who want to enforce political compliance and prevent rebellion. More mildly, marketing companies could use these techniques to modulate customers’ preferences and attitudes towards their products.

Just like the previous two rights, it is a matter of discussion whether the right to psychological continuity should be considered absolute or relative. It could be argued that some neurotechnologically-induced personality changes could be tolerated with regard to persistent violent offenders (for instance, serial rapists, killers and pedophiles). The need to protect the public from potentially dangerous individuals who are very likely to reoffend if released would justify such measures. This would even be a good alternative for those individuals themselves, who could avoid in this way spending their whole lives in prison. However, extreme caution and broad public discussion is imperative before authorizing such intentional intrusions into people’s personality.

Conclusions

The volume and variety of neurotechnology applications is rapidly increasing inside and outside the clinical and research setting. The ubiquitous distribution of cheaper, scalable and easy-to-use neuroapplications has the potential of opening unprecedented opportunities at the brain-machine interface and making neurotechnology intricately embedded in our everyday life. While this technological trend may generate immense advantage for society at large in terms of clinical benefit, prevention, self-quantification, bias-reduction, personalized technology use, marketing analysis, military dominance, national security and even judicial accuracy, yet its implications for ethics and the law remain largely unexplored. We argue that in the light of the disruptive change that neurotechnology is determining in the digital ecosystem, the normative terrain should be urgently prepared to prevent misuse or unintended negative consequences. In addition, given the fundamental character of the neurocognitive dimension, we argue that such normative response should not exclusively focus on tort law but also on foundational issues at the level of human right law.

In this context, we have suggested that emerging trends in neurotechnology are eliciting coordinate amendments to the current human right framework which will require either a reconceptualization of existing human rights or even the creation of new neuro-specific rights. In particular, we have argued that emerging collateral risks associated with the widespread use of pervasive neurotechnology such as malicious brain-hacking as well as hazardous uses of medical neurotechnology may require a reconceptualization of the right to mental integrity. In fact, although mental integrity is protected by the EU Charter of Fundamental Rights (Article 3), this right is conceptualized as a right to accessing and
protecting mental health and is complementary to the right to physical integrity. We suggest that in response to emerging neurotechnology possibilities, the right to mental integrity should not exclusively guarantee protection from mental illness or traumatic injury but also from unauthorized intrusions into a person’s mental wellbeing performed through the use of neurotechnology, especially if such intrusions result in physical or mental harm to the neurotechnology user.

In addition to such reconceptualization, we have argued that the creation of neuro-specific rights may be required as a coping strategy against possible misuses of neurotechnology as well as a form of protection of fundamental liberties associated with individual decision-making in the context of neurotechnology use. With this respect, we have endorsed the recognition of a negative right to cognitive liberty as a right for the protection of individuals from the coercive and unconsented use of such technologies. In addition, as a complementary solution, we have proposed the recognition of two additional neuro-specific rights: the right to mental privacy and the right to psychological continuity. The right to mental privacy is a neuro-specific privacy right which protects private or sensitive information in a person’s mind from unauthorized collection, storage, use, or even deletion — in digital form or otherwise. In contrast to existing privacy rights, the right to mental privacy stands to protect information prior to any extra-cranial externalization (e.g. in verbal or printed format) as well as the generator of such information (a person’s neural processing). As such, it protects a person’s mental dimension as the ultimate domain of information privacy in the digital ecosystem. In coordination with that, the right to psychological continuity will protect the mental substrates of personal identity from unconscious and unconsented alteration by third parties through the use of invasive or non-invasive neurotechnology.

All these proposed neuro-focused rights are mutually linked and stand in an intimate family relationship. Being the substrate of all other freedoms, cognitive liberty in its positive sense is a prerequisite of all other neuro-focused rights. As such, it is to mental privacy, mental integrity and psychological continuity in a very similar relation as freedom of thought is to privacy, integrity and identity rights. However, in its negative sense of protection from coercive use, cognitive liberty can only partly account for unintended uses of emerging neurotechnology. In fact, illicit intrusions into a person’s mental privacy may not necessarily involve coercion, as they could be performed under the threshold of a persons’ conscious experience. The same goes for actions involving harm to a person’s mental life or unauthorized modifications of a person’s psychological continuity, which are also facilitated by the ability of emerging neurotechnologies to intervene into a person’s neural processing in absence of the person’s awareness.

This proposal of neuro-specific human rights in response to emerging advancements in neurotechnology is consistent with and a logical continuation of the proposal of developing genetic-specific human rights in response to advancements in genetics and genomics as set out by the Universal Declaration on the Human Genome and Human Rights (UDHGHR) and the International Declaration on Human Genetic Data (IDHGGD).

Extensive future debate is required to test the normative solidity of this proposed expansion of the human right framework to the neurotechnology dimension. In parallel, future research is required to investigate the implications of such proposed human rights on other levels of law such as international humanitarian law, criminal law, tort law, property law and consumer law. Ideally, this debate should benefit from the active and cross-disciplinary participation of legal experts, neuroscientists, technology developers, neuroethicists and regulation bodies.
Endnotes

1For example, the website The Brain Stimulator offers a wide assortment of affordable tDCS devices, with prices ranging between 60$ and 200$. See: https://thebrainstimulator.net/shop/


4The expression “reasonable expectation of privacy” was coined by the US Supreme Court in 1967 to distinguish legitimate police searches and seizures from unreasonable ones in the light of the Fourth Amendment that protects privacy rights.

5In the exercise of his rights and freedoms, everyone shall be subject only to such limitations as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others and of meeting the just requirements of morality, public order and the general welfare in a democratic society” (UDHR, Article 29.2).

6American Convention on Human Rights, art. 8(2)(g): “Every person accused of a criminal offense has the right (...) not to be compelled to be a witness against himself or to plead guilty”; Rome Statute of the International Criminal Court, art. 55(1)(a): “In respect of an investigation under this Statute, a person: (a) Shall not be compelled to incriminate himself or herself or to confess guilt”. Paragraph 2(b) of the same Article 55 adds that the person suspected of having committed a crime has the right “to remain silent, without such silence being a consideration in the determination of guilt or innocence”.

7Funke v. France, ECtHR 1993, A n° 256-A; John Murray v. United Kingdom, ECtHR 1996-I.


9Saunders v. United Kingdom, ECtHR 1996-VI, para 69.

10The expression “Manchurian candidate” refers to “a person who is (or is believed to be) brainwashed into becoming a subversive agent, especially an assassin” (Oxford Dictionary). The expression was popularized by the 1962 film The Manchurian Candidate, adapted from the 1959 novel of the same name by Richard Condon.


Authors’ contribution
MI and RA jointly developed the conceptual structure, logical articulation and equally contributed to the draft of the manuscript. Each author revised the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Competing interests
None of the authors have any competing interests in the manuscript.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Institute for Biomedical Ethics, University of Basel, Bernoullistrasse 28, 4056 Basel, Switzerland. 2School of Law, University of Zurich, Zurich, Switzerland.

Received: 20 October 2016 Accepted: 20 March 2017
Published online: 26 April 2017

References


Panel 1: The Future of Neuroscience and its Moral (Ethical) Implications

Gray Matters: Topics at the Intersection of Neuroscience, Ethics, and Society

The Presidential Commission for the Study of Bioethical Issues, which produced this report, is an advisory panel of the nation’s leaders in medicine, science, ethics, religion, law, and engineering. The purpose of this report is to consider the ethical and societal implications of neuroscience research and its applications, which will allow us to better understand the human brain. The report maintains that this research will result in the development of more effective diagnostic tools, treatments, preventions, and cures for neurological disorders and psychiatric conditions that affect tens of millions of individuals in the United States alone.

Neuroethics in the Age of Brain Projects

This article explores two large brain projects, the US BRAIN Initiative and the European Union Human Brain Project, and examines the neuroethical implications of the two projects. The article gives an overview of both projects and emphasizes the need for neuroethics as our knowledge of the brain continues to grow.

Panel 2: The Ethics of Trauma, Aging, and Neuroscientific (Brain) Irregularities (Deviations)

Brain Equality: Legal Implications of Neurodiversity in a Comparative Perspective

This article examines the concept of neurodiversity and its impact on the legal system. Neurodiversity refers to “a broad spectrum of cognitive, linguistic, and learning functions, and early-onset neurobiological conditions that impair the individual's capacity for social understanding, social inter-action, learning, and pragmatic and semantic communication.” The article argues that the consideration of neurodiversity could have vast constitutional implications as well as implications concerning legal systems in both the US and in Europe.
**Panel 3: The Law and Ethics of Neuroscientific Development**

**Scanning for Justice: Using Neuroscience to Create a More Inclusive Legal System**

This article examines how neurolaw, through its study of human behavior, can inform law, legislation, and policy. For example, neurolaw could have a great impact on civil and human rights by making it easier for plaintiffs to bring claims that historically have been difficult to prove. The article argues that the use of neuroscience generates a fundamentally fairer legal system, which will be able to detect previously hidden or latent injuries, especially those in marginalized communities.

**Towards New Human Rights in the Age of Neuroscience and Neurotechnology**

This article explores the human rights implications that need to be addressed as neuroscience and neurotechnology allow for an unprecedented amount of information to be accessed, collected, shared and manipulated. It analyzes the relationship between neuroscience and human rights, and then identifies new rights which may become relevant as neuroscience becomes more advanced.
2016

A Glimpse Inside the Brain’s Black Box: Understanding the Role of Neuroscience in Criminal Sentencing

Bernice B. Donald
U.S. Court of Appeals for the Sixth Circuit

Erica Bakies

Recommended Citation
Available at: http://ir.lawnet.fordham.edu/flr/vol85/iss2/5

This Symposium is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Fordham Law Review by an authorized editor of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact tmelnick@law.fordham.edu.
INTRODUCTION

It is not a secret: size matters. And where it matters most is within the most complex structure in the universe—the brain, a mass of gray and white matter that controls an extraordinary number of functions and processes that allow us to walk, talk, breathe, reason, feel emotions, and perceive and experience the world around us. While we have made great strides in studying this three-pound ball of cells, it still mostly remains a mystery beyond our grasp of comprehension. But what little we do know has led to great developments in the legal community and especially in the criminal justice system. This Article focuses on the utilization of neuroscience and its developing technology in the courtroom, particularly at the sentencing phase of trial.

While the brain encompasses a wide variety of fields of study, neuroscience offers specific and tangible insight into brain underdevelopment and brain injuries. For example, neuroscience demonstrates that what our childhood was like—whether good, bad, or in between—greatly impacts the full development of this vital organ. Studies show that exposure to stress and instability actually prevents the brain from fully developing. In other words, the brain remains small and those processes it controls immature. Children exposed to trauma face a number of disorders, including “depression, attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), conduct disorder, anxiety disorders, eating disorders, sleep disorders, communication disorders, separation anxiety disorder, and/or reactive attachment disorder,” to name a few.1

* Judge, U.S. Court of Appeals for the Sixth Circuit. The authors would like to recognize Kara Bidstrup, Michael Poupore, and Lyle Gruby for their contributions. This Article is part of a symposium entitled Criminal Behavior and the Brain: When Law and Neuroscience Collide held at Fordham University School of Law. For an overview of the symposium, see Deborah W. Denno, Foreword: Criminal Behavior and the Brain: When Law and Neuroscience Collide, 85 FORDHAM L. REV. 399 (2016).

** Law Clerk to the Honorable Bernice B. Donald.

1. Alexandra Cook et al., Complex Trauma in Children & Adolescents, 21 FOCAL POINT 4, 4 (2007).
Because the effects of childhood trauma stem from one’s surroundings, children who live in inner cities that are plagued with violence and racial tension often experience childhood trauma and the diagnoses that follow it. Notably, the same types of experiences are present in children who are exposed to the welfare system. While most people are familiar with posttraumatic stress disorder (PTSD), children who are constantly exposed to trauma and dangerous situations, like those in dangerous inner-city neighborhoods, face a far more compounded scenario: complex trauma. Complex trauma is “the dual problem of children’s exposure to multiple traumatic events and the impact of this exposure on immediate and long-term outcomes.” Those outcomes include “psychiatric and addictive disorders, chronic medical illness, and legal, vocational, and family problems.” In short, neuroscience can identify both childhood trauma and its lasting impression on an individual as they become an adult.

Now that current neuroscience technology has the ability to demonstrate how exposure to childhood trauma affects an individual’s brain, the next question is how this science and its conclusions in the courtroom can be effectively utilized. This question becomes very apparent in the context of sentencing, where a judge may consider a wide range of factors in determining an appropriate sentence for those defendants standing before her. Without disregarding the criminal justice system’s ability to hold those accountable for their actions, neuroscience can be utilized to demonstrate that certain actions may actually be the result of developmental problems associated with the brain, like the effects of complex trauma on children. A judge may also use neuroscience to combat her implicit biases, which have ways of manifesting themselves in the courtroom and therefore need to be explicitly acknowledged. Neuroscience can offer additional insight into a defendant’s thought process and accordingly provide a means for the judge to address and correct those biases.

This Article begins by discussing what neuroscience and the smaller associated field of study, neuropsychology, are and what they can tell us about an individual. It then recounts a brief history of sentencing in the United States. Additionally, it expounds on how the legal system currently utilizes neuroscience in the courts, noting specifically the ways in which neuroscience can be presented during the sentencing phase of trial. Finally, it discusses the use of neuroscience as a mitigating factor during sentencing and how judges can use neuroscience to combat their implicit biases.

I. BACKGROUND

In conjunction with the National Institute of Health, President Barack Obama launched the Brain Research through Advancing Innovative
Neurotechnologies (BRAIN) Initiative, which focuses on revolutionizing our understanding of the human brain. The goal is to fill major gaps in our current knowledge and provide unprecedented opportunities for exploring exactly how the brain enables the human body to record, process, utilize, store, and retrieve vast quantities of information, all at the speed of thought. Recent developments concerning the brain have been historic and far reaching.

A. The Science

Unsurprisingly, this expansion in funding and focus has contributed to the many fields of study concerning the brain. Relevant here, neuroscientists are not only concerned with the normal functioning of the brain in conjunction with the rest of the nervous system but also with the effects of neurological, psychiatric, and developmental disorders on people’s actions. Neuroscience is a branch of the life sciences that addresses the anatomy, physiology, biochemistry, or molecular biology of nerves and nervous tissue within the brain, specifically in relation to behavior and learning. More generally, neuroscience is the study of how processes function within the brain.

Another key area of study that increases our understanding of the complexity of the brain and its effects on behavior is neuropsychology. Neuropsychology adds to the picture outlined by neuroscience, as it is concerned with the integration of psychological observations of behavior and the mind with neurological observations of the brain and nervous system. This is because “[s]tandard neuroimaging is neither specific nor sensitive enough to detect the damage done to the brain.” Basically,

8. Id. (explaining that by “accelerating the development and application of innovative technologies, researchers will be able to produce a revolutionary new dynamic picture of the brain that . . . shows how individual [brain] cells and complex neural circuits interact . . . at the speed of thought”).
12. Bruce H. Stern, Neuropsychology & Traumatic Brain Injury, TRIAL, Oct. 2015, at 48, 49 (“Advanced neuroimaging, such as a PET scan or diffusion tensor imaging, provides evidence about how the brain is functioning . . . .”)).
neuropsychology objectively analyzes how the mind works in connection with the brain.13

Neuroscience and neuropsychology typically involve comprehensive and extensive evaluations. These evaluations have been incorporated into the legal field most notably as “neuroscience evidence.”14 Two of the most common categories of these tests are (1) neuroimaging, or “imaging tests,” which are generated by computer images of a human brain, and (2) neuropsychological exams, or “non-imaging tests,” which are based on tests administered by a medical professional to an individual for the purpose of gaining insight into how that person’s brain operates.”15

Neuroimaging now allows neurologists to analyze the structural and functional aspects of the brain. Structural neuroimaging involves magnetic resonance imaging (MRI) and computed tomography (CT) scans.17 These images demonstrate “the brain’s architecture.”18 Similarly, scans such as the electroencephalography (EEG), positron emission tomography (PET) scans, and functional magnetic resonance imaging (fMRI) display visual images of how the brain works at a particular moment in time.19

Neuropsychological exams are more than just scans of the brain. A neuropsychological evaluation is a “comprehensive, objective assessment of a wide range of cognitive, adaptive, and emotional behaviors that reflect the adequacy or inadequacy of higher brain functions.”20 In other words, neuropsychological testing measures a person’s brain function compared to the normal population in a variety of different areas, including education, standardized test scores, and work history.21 This variety of tests—also known as a “battery”—requires access to the subject’s school records, medical records, and employment records.22 Other relevant background information may include the subject’s social and family history.23 Considering all of these factors, a neuropsychologist can then determine a baseline of brain function for a particular individual, generally before a particular event occurs, such as the committing of a crime.24 For example, after conducting these tests, neuropsychologists are able to better understand and interpret the consequences of childhood neglect and its

13. See id.
15. Id.
17. See id.
18. Id. (quoting Joshua Greene & Jonathan Cohen, For the Law, Neuroscience Changes Nothing and Everything, 359 Phil. Transactions Royal Soc’y London B 1775, 1775 (2004)).
21. See id. at 49–50.
22. See id. at 49.
23. See id.
24. See id.
effects on brain development, particularly when it comes to explaining how those individuals ended up in the criminal justice system.\textsuperscript{25}

Because “human behavior is the very currency in which law”—and especially criminal law—deals, there is an unending need for an improved understanding of how and why particular people behave the way that they do.\textsuperscript{26} Dubbed “neurolaw,” this “neuroscience revolution” has gained the attention of legal thinkers and is poised to be the catalyst for significant changes in not only the criminal justice system but the legal field generally.\textsuperscript{27}

The development of neuroscience includes recent momentous breakthroughs, especially in how certain types of experiences during childhood can drastically affect the rest of a child’s life.\textsuperscript{28} In utero and during the first four years of life, a child’s rapidly developing brain organizes to reflect the child’s environment.\textsuperscript{29} By the age of four, a child’s brain is 90 percent of its adult size.\textsuperscript{30} Accordingly, a child who is exposed to trauma early on in her life organizes her brain around instability and chaos, which is extremely debilitating.\textsuperscript{31} For example, receiving “proper nutrition and stimulation during the first three years of life” is critical “for the brain to develop the crucial neurological networks that are foundational to the functioning of an individual.”\textsuperscript{32} Because of neuroscience, we now know that having these types of experiences at an early age can lead to “permanent and irreversible consequences,” especially in the “physical, cognitive, emotional, and social domains.”\textsuperscript{33} It bears noting that these consequences can weigh heavily not only on the child exposed to such circumstances but also on society itself, a weight that often goes unrecognized.\textsuperscript{34}

One of the most interesting aspects of neuroscience is that it can show us actual physical changes in response to childhood trauma.\textsuperscript{35} Studies demonstrate that there are differences in the volume of an adult’s prefrontal

\begin{itemize}
\item \textsuperscript{25} Janet Weinstein & Ricardo Weinstein, Before It’s Too Late: Neuropsychological Consequences of Child Neglect and Their Implications for Law and Social Policy, 33 U. MICH. J.L. REFORM 561, 562 (2000).
\item \textsuperscript{26} Owen D. Jones & Timothy H. Goldsmith, Law and Behavioral Biology, 105 COLUM. L. REV. 405, 407 (2005).
\item \textsuperscript{28} See Weinstein & Weinstein, supra note 25, at 562 (“Recent developments in the neurosciences have led to dramatic breakthroughs in the area of brain development and [especially] the understanding of consequences of [childhood] neglect.”).
\item \textsuperscript{29} See Bruce D. Perry, Maltreatment and the Developing Child: How Early Childhood Experience Shapes Child and Culture, MARGARET MCCAIN LECTURE SERIES 2 (Sept. 23, 2004), http://www.flcc.on.ca/mccain/perry.pdf [https://perma.cc/PA8T-EYJG].
\item \textsuperscript{30} See id.
\item \textsuperscript{31} See id.
\item \textsuperscript{32} Weinstein & Weinstein, supra note 25, at 561.
\item \textsuperscript{33} Id. at 595.
\item \textsuperscript{34} See id.
\end{itemize}
cortex depending on whether that individual experienced trauma as child or whether they had a nurturing childhood. Studies also indicate that maltreated children exhibited “higher rates of adult psychopathology and a greater likelihood of engaging in maladaptive and socially disruptive courses of conduct as adults (such as engaging in substance abuse or violating criminal law).” These findings are significant in establishing the connection between childhood trauma and future offenses because the prefrontal cortex is involved in a vast number of functions, such as “executive functionality (for example, planning and controlling behavioral responses, problem-solving, and sustaining mental productivity), attention focusing, working and delayed memory, emotional regulation, and responses to stress.”

A newer development in the cross section of childhood trauma, neuropsychology, and neuroscience is analysis of “complex trauma” and its long-term effects on children. When a child is exposed to any threat, her brain will activate a set of adaptive responses designed to help her survive. “Complex trauma” occurs when a child has been exposed to multiple traumatic events throughout her early life. When a child experiences repetitive activation of the stress response systems, her baseline state of arousal is altered. Thus, even when there is no external threat or demand, she is in a psychological state of alarm, commonly known as “fight or flight.” Although PTSD is similar to the fight-or-flight state of mind that exists with complex trauma, it does not capture the full range of developmental difficulties that traumatized children experience as a result of exposure to repeated traumatic incidents. For instance, when a stressor arises, which could be as simple as an argument with a peer or a demanding school task, a traumatized child’s emotions may rapidly escalate to a state of fear.

Further, when in a state of calm, a person can use the higher, more complex parts of the brain to process and act on information; in contrast, in a state of fear, a person is only able to access the lower, more primitive parts of the brain. An increase in threat level corresponds to “less thoughtful and . . . more reactive” responses. Further, “[a]ctions in this state may be governed by emotional and reactive thinking styles.” Because a traumatized child’s baseline state of arousal is constantly altered,
she is unable to learn from normal “social, emotional, and other life
experiences.”

Complex trauma generally manifests itself in children who experience
abuse or neglect, but it can also appear in children who have witnessed
domestic violence, ethnic cleansing, or war. The consequences of
complex trauma on a child are devastating for both her and her future. That
child can experience, among other things, trouble with “accurate
identification of internal emotional experiences” and interference “with the
formation of a secure attachment bond between a child and her caregiver,”
which influences the child’s future relationships and social skills. The
child can carry all of these problems forward to adulthood, where she must
deal with them as a functional and participating member of society.

Children who have complex trauma or who have survived extreme
neglect are among the “hundreds of millions of people around the world
living with mental disorders.” The World Health Organization’s (WHO)
Mental Health Action Plan 2013–2020 emphasizes that, depending on the
local context, certain portions of the population are more susceptible to
mental health issues than others. Members of households living in
poverty and infants or children exposed to maltreatment and neglect are
included in those vulnerable categories. The WHO contextualizes these
matters as more than public health issues; it characterizes them as
developmental issues. If these issues are not directly addressed, children
will continue to suffer from developmental impairments, perhaps by
committing crimes or by being unable to positively contribute to society.
The number of those suffering continues to increase, and if they are
prevented from becoming productive members of society, the global
economy will suffer as well.

48. Id.
49. See id. at 2; see also Cook et al., supra note 1, at 1.
50. Cook et al., supra note 1, at 4–5.
51. Making Mental Health a Global Development Priority, MHGAP NEWSL. 1 (May
2016) [hereinafter MHGAP NEWSLETTER], http://www.who.int/mental_health/mhgap/
newsletter_may_2016.pdf?ua=1 [https://perma.cc/W86G-ECP7].
52. WORLD HEALTH ORG., MENTAL HEALTH ACTION PLAN 2013–2020, at 7 (2013),
http://www.who.int/mental_health/action_plan_2013/bw_version.pdf [https://perma.cc/C7
Y6-MSHR].
53. See id.
54. MHGAP NEWSLETTER, supra note 51; see also Out of the Shadows: Making Mental
16, 2016) [https://perma.cc/426B-ASXX].
55. MHGAP NEWSLETTER, supra note 51.
56. See id. (noting that the World Bank, which is composed of ministers of finance and
development agencies, decided to join the WHO in its efforts to move mental health into the
mainstream developmental agenda).
B. Sentencing in the United States

Imprisonment in the United States and elsewhere was uncommon prior to the eighteenth century and was not used as a primary form of punishment until around the American Revolution.\(^57\) Punishment prior to prisons often focused on retaliation and vengeance.\(^58\) The Quakers, as pacifists, were against capital punishment and, accordingly, crusaded for reform by suggesting replacing traditional punishment with confinement and labor in prisons.\(^59\) This stance abruptly changed during the mid-twentieth century, and instead of confinement and labor, punishment focused on rehabilitation.\(^60\) About thirty to forty years ago, sentencing was again reformed.\(^61\) Pivoting from a focus on rehabilitation during the 1970s and 1980s, public policy instead began applying severe penalties associated with the penological goals of deterrence and incapacitation.\(^62\)

The pivot in the 1970s and 1980s corresponded with an increase in street drugs, such as crack cocaine.\(^63\) It was this rise in drug use that led to the marked “War on Drugs.”\(^64\) The War on Drugs drastically increased the incarceration rate in numerous ways:

1. the direct incarceration of drug offenders,
2. the re-incarceration of all types of offenders due to drug-related parole violations,
3. the impact of drug incarcerations on prison admissions instead of prison populations,
4. the extent to which prior drug offenses trigger repeat-offender enhancement, even for non-drug crimes, and
5. the effects of large-scale drug arrests and incarcerations on neighborhood social cohesion, and the connections between social stability and incarceration.\(^65\)

In all, drug convictions increased tenfold between 1980 and 1996.\(^66\)

The War on Drugs also had an unintended but formidable effect on individuals with mental health issues who were susceptible to coming into


\(^{60}\) See *id.* at 849.


\(^{62}\) See *id.*

\(^{63}\) See Paul Butler, *Retribution, for Liberals*, 46 UCLA L. REV. 1873, 1884 (1999). But see Doris Marie Provine, *Race and Inequality in the War on Drugs*, 7 ANN. REV. L. & SOC. SCI. 41, 4849 (2011) (“The war on drugs is thus distinctive from, though not incompatible with, the hardening of attitudes toward crime and punishment that began to take form in the 1970s.”).

\(^{64}\) See Butler, *supra* note 63, at 48–49.


contact with the criminal justice system.\textsuperscript{67} Prior to the War on Drugs, the United States began deinstitutionalizing psychiatric facilities with the hope of returning those individuals to nursing homes, assisted living facilities, and home care.\textsuperscript{68} When the last penny of government funding was spent, these asserted alternatives were underutilized and many of the mentally ill ended up on the streets where they self-medicated their illnesses.\textsuperscript{69} The War on Drugs started soon thereafter, in the 1970s.\textsuperscript{70} The culmination of these two trends led to a jump in the percentage of inmates with serious mental illness as well as a large decrease in the number of individuals hospitalized for serious mental illness.\textsuperscript{71}

The positive correlation between mental health issues and addiction is embodied in the concept of “dual diagnosis,” a situation where a person has both a mood disorder and an addiction problem; thus, the War on Drugs’s high incarceration periods for drug convictions disproportionately impacted those individuals living with mental health issues.\textsuperscript{72} The National Alliance on Mental Illness estimates that “[a]bout a third of all people experiencing mental illnesses and about half of people living with severe mental illnesses also experience substance abuse.”\textsuperscript{73}

Prior to 1984, judges had almost unlimited discretion when it came to sentencing.\textsuperscript{74} One commentator noted that the system was one that had “the absence of rational ordering,” with various sentences applied “arbitrar[ily] and discriminator[ily].”\textsuperscript{75} Accordingly, in 1984, Congress attempted to stem judges’ broad sentencing discretion by replacing it with a more uniform application of prescribed sentences.\textsuperscript{76} As part of the Comprehensive Crime Control Act, Congress passed the Sentencing Reform Act.\textsuperscript{77} That statute, in turn, established the U.S. Sentencing Commission (“the Sentencing Commission” or “the Commission”), an agency of the judicial branch that became responsible for developing uniform guidelines for sentencing.\textsuperscript{78} The Sentencing Commission kept

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{67} See Terry A. Kupers, A Community Mental Health Model in Corrections, 26 STAN. L. & POL’Y REV. 119, 123 (2015) (noting that the War on Drugs “captured many individuals with serious mental illness in its dragnet”).
\item \textsuperscript{68} Anisha Lewis, Incarceration and Mental Health, CTR. PRISONER HEALTH & HUM. RTS., http://www.prisonerhealth.org/educational-resources/factsheets-2/incarceration-and-mental-health (last visited Oct. 16, 2016) [https://perma.cc/4P72-7HQD].
\item \textsuperscript{69} Id.
\item \textsuperscript{70} Id.
\item \textsuperscript{71} Dual Diagnosis, NAT’L ALLIANCE ON MENTAL ILLNESS, http://www.nami.org/Learn-More/Mental-Health-Conditions/Related-Conditions/Dual-Diagnosis (last visited Oct. 16, 2016) [https://perma.cc/3KLM-CCKN].
\item \textsuperscript{72} Id.
\item \textsuperscript{73} Id.
\item \textsuperscript{74} MARVIN E. FRANKEL, CRIMINAL SENTENCES: LAW WITHOUT ORDER 49 (1973).
\item \textsuperscript{75} Id.
\item \textsuperscript{78} See Mistretta v. United States, 488 U.S. 361, 363–70 (1989) (discussing the background, purpose, and operation of the Sentencing Reform Act and the Commission).
\end{enumerate}
\end{footnotesize}
some of the more traditional goals in mind when developing guidelines: punishment, deterrence, incapacitation, and rehabilitation. At the same time, the Commission also focused on “minimizing disparity in sentencing, and accounting for ‘advancement in the knowledge of human behavior as it relates to the criminal justice system.’” Thus, the Commission promulgated the U.S. Sentencing Guidelines (“the Sentencing Guidelines” or “the Guidelines”), which sought to correct the discriminatory application of sentences and provide for more predictable sentences. These mandatory Guidelines were promulgated in 1987, ending an era in which a trial judge’s prescribed sentence was “virtually unquestioned.” It was not until 2005 that the U.S. Supreme Court ruled that the Sentencing Guidelines, which Congress had made mandatory for judges to follow, were unconstitutional. The Court held that requiring judges to sentence within a set time period violated defendants’ Sixth Amendment rights. As a result, the Guidelines became “effectively advisory.” This brought a sea of change to the federal appellate courts. Some took the “Booker maximalism” stance, the view that “the Guidelines no longer had a privileged place in sentencing.” Others applied a “Booker minimalism” approach, which gave the Guidelines more weight than other factors, arguing that “they had a special role in promoting sentencing uniformity,” and they “accounted for the other § 3553(a) factors.”

The effects of the Sentencing Guidelines—both when they were mandatory and now that they are advisory—and the War on Drugs are not all positive and their combined effect on society, some argue, has been grave. Large differences in the length of sentences still exist on the basis of race, gender, education, income, and citizenship, despite the Guidelines’ command that these characteristics not affect the sentence length. For

79. Id.
81. Id.
84. See Booker, 543 U.S. at 245.
85. Id.
87. Id. at 836.
88. See, e.g., Crystal S. Yang, Free at Last?: Judicial Discretion and Racial Disparities in Federal Sentencing, 44 J. LEGAL STUD. 75, 76 (2015) (“While the guidelines reduced inter judge sentencing disparities in their early years, . . . many criticized them for being rigid . . . and for shifting power to prosecutors in their charging and plea-bargaining decisions.”).
example, racial disparities in sentencing continue to exist, and some studies suggest that they have actually increased since *Booker*. African Americans account for approximately 12 percent of the population of the United States, but almost 40 percent of those incarcerated are African American. There are more African American men in prison today than there were African Americans slaves in 1850. Lower-income offenders are less likely to receive downward departures and more likely to receive upward departures. Evidence suggests that even when judges have awarded departures from the recommended sentences to lower-earning offenders, these departures are typically only small reductions in sentencing. Because the Guidelines allow judges to reduce or increase the recommended sentence, over half of the unaccounted for differences in sentences are generated by departures from the Guidelines, rather than from sentencing within the Guidelines. While the Guidelines are certainly Congress’s attempt to combat these disparities, they clearly still exist, and there is at least the suggestion that implicit bias in sentencing and prosecution play a role in maintaining that disparity today.

Legal commentators suggested that the Guidelines had a similar effect on the mentally ill. The Guidelines provided for a downward departure from the calculated sentencing Guidelines range for individuals with a mental illness. In considering whether to apply this downward departure, the Guidelines instruct judges to consider “if such [mental and emotional] conditions, individually or in combination with other offender characteristics, are present to an unusual degree and distinguish the case from the typical cases covered by the guidelines.” The Guidelines are careful to focus on mental and emotional conditions, which were previously deemed irrelevant to determining whether a downward departure was warranted in a particular situation. In addition to those considerations,
the Guidelines account for diminished capacity in section 5K2.13, which also allows for downward departures from the recommended Guidelines sentencing range.

Despite these lofty goals and far-reaching considerations for mental health, pundits suggest that these outcomes have not been realized. Instead of providing for further consideration of mental illness, legal commentators have observed, “Booker’s main effect may have been to create a second pathway for judges to impose above-Guidelines sentences.” Some say this reaction can be traced back to one of the main impetuses for establishing the Guidelines: John Hinckley’s acquittal and the subsequent public distaste for the insanity defense. The War on Drugs also may have played a role here, as “the Guidelines were crafted to ensure that drug dependence, which is perhaps most reasonably viewed as mental illness, would not act to mitigate sentences.”

Even though there may be issues with the Sentencing Guidelines, they are the current method by which the judiciary bases sentencing decisions. Utilizing neuroscience in this space requires consideration and knowledge of the Guidelines’ advantages as well as their disadvantages.

C. Current Utilization of Neuroscience in the Criminal Justice System

Lawyers recently have begun utilizing neuroscience in the courtroom, and, almost more notably, courts have embraced it. For instance, in 2005, the Supreme Court considered neuroscientific theories of child development to support its reasoning in prohibiting the death penalty for older juveniles. More recently, in a 2011 case, Brown v. Entertainment Merchants Ass’n, Justice Breyer wrote a dissent that relied heavily on neuroscience research demonstrating a correlation between virtual violence in video games and aggressive tendencies of those children who played them. While “law and neuroscience” is useful as a general descriptive phrase, it is too vague to be applicable in particular research and applied contexts. In practice, it is particular aspects of law that may be affected by particular types of neuroscience research.

101. See id. § 5K2.13.
102. See id.
103. The Law of Mental Illness, supra note 97, at 1138.
104. Id. at 1135.
105. Id. at 1136.
108. Id. at 850–56 (Breyer, J., dissenting).
1. At Trial

Neuroscience has slowly crept into the trial phase of court cases, in both the criminal and civil dockets.\textsuperscript{110} In the criminal docket, it has been introduced in all three phases of trial: preliminary determinations of competency, the guilt phase, and the sentencing phase.\textsuperscript{111} Most notably, neuroscience has been utilized to negate an ability to form the mens rea necessary for premeditation and deliberation and to bolster a defense for not guilty by reason of insanity.\textsuperscript{112} Even though these types of cases are few and far between, and the science is still in its infancy, they demonstrate the roles that neuroscience could one day play on a grander scale.

In the civil docket, neuroscience has proved itself to be equally beneficial. For example, in \textit{P.P. v. Compton Unified School District},\textsuperscript{113} the plaintiffs, students in the Compton Unified School District in Los Angeles, California, utilized neuroscience to bolster their complaint.\textsuperscript{114} The district court relied on the scientific evidence to deny a motion dismiss for failure to state a claim.\textsuperscript{115} The students’ claim was that their exposure to childhood trauma\textsuperscript{116} and their likely diagnosis of complex trauma “impair[ed] their ability to perform activities essential to education—including, but not limited to, learning, thinking, reading, and concentrating,” resulting in a violation of the Individuals with Disabilities Education Act (IDEA).\textsuperscript{117} All sorts of claims have utilized neuroscience in the civil arena: personal injury, medical malpractice, and toxic exposure cases, to name a few.\textsuperscript{118} In short, neuroscience’s applications are widespread, and lawyers should consider whether it could play a crucial role in litigation.

2. At Sentencing

Neuroscience can also play a unique role in sentencing, although the scope of that role is still somewhat vague. There is, of course, the issue of admissibility, which is currently being debated in state courts. Like other experts presented during sentencing, either Federal Rule of Evidence 702 or the associated state rule usually applies to expert testimony provided by neuropsychologists regarding diagnosis and causation.\textsuperscript{119} While most states

\textsuperscript{111} See id. at 341.
\textsuperscript{112} See id. at 341–42.
\textsuperscript{113} 135 F. Supp. 3d 1126 (C.D. Cal. 2015).
\textsuperscript{114} Id. at 1147–49.
\textsuperscript{115} Id.
\textsuperscript{116} Some representative examples include a child being “repeatedly physically and sexually abused by his mother’s boyfriends,” witnessing the physical abuse of family members, watching a best friend be shot and killed, being “stabbed with a knife while trying to protect a friend,” and being “sexually assaulted on the bus on her way home from school,” among other horrible experiences. Id. at 1130.
\textsuperscript{117} Id. at 1131.
\textsuperscript{118} See Compton, supra note 110, at 341.
\textsuperscript{119} See Stern, supra note 12, at 50.
accept neuropsychological testimony from experts, a minority of states—such as Florida, Georgia, and Virginia—prohibit it, instead concluding that only medical experts, such as physicians, are qualified to testify on diagnosis, causation, and prognosis.\footnote{120}

At the federal level, the Sixth Circuit has weighed in on the debate. In \textit{Fautenberry v. Mitchell},\footnote{121} the dissenting opinion stated that a neuropsychological examination is “the most effective means possible of determining whether [the defendant] had a brain impairment” stemming from her unstable family environment, emotionally abusive upbringing, and undocumented physical abuse.\footnote{122} Although the defendant declined to submit to any neurological exams in that particular case, the dissent went on to acknowledge that a neuropsychologist could presumably have discovered, verified, and revealed such brain damage.\footnote{123} Further, the dissenting opinion argued that this evidence could have been admitted to aid the defendant’s case as mitigating evidence during the sentencing stage of trial.\footnote{124}

Perhaps unsurprisingly, neuroscience is mostly utilized in very serious cases, generally where the defendant is facing a death sentence, life imprisonment, or a substantially long term of imprisonment.\footnote{125} Specifically with respect to the death penalty, the Supreme Court has recognized the opportunity to consider offenders’ past life experiences and other evidence.\footnote{126} It has concluded the following:

\begin{quote}
[T]he Eighth and Fourteenth Amendments require that the sentencer, in all but the rarest kind of capital case, not be precluded from considering, as a mitigating factor, any aspect of a defendant’s character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death.\footnote{127}
\end{quote}

The Court described these types of considerations in death penalty cases “far more important than in noncapital cases.”\footnote{128}

Jurisprudence and legal commentators have noted that considering mitigating circumstances is critical because they suggest that the defendant is not fully culpable for the crime charged, and therefore, the defendant is worthy of a lower sentence than the average person who has no neurological issues.\footnote{129} A prime example of this logic is reflected in the Supreme Court’s 2002 case \textit{Atkins v. Virginia}.\footnote{130} There, the Court held that

\begin{itemize}
\item[120.] See id.
\item[121.] 515 F.3d 614 (6th Cir. 2008).
\item[122.] \textit{Id.} at 625 (Moore, J., dissenting).
\item[123.] \textit{Id.}
\item[124.] \textit{Id.} at 645.
\item[125.] See Denno, \textit{supra} note 10, at 502 (“In sum, my analysis indicates that neuroscience evidence is typically used in cases where defendants face the death penalty, a life sentence, or a substantial prison sentence.”).
\item[126.] See \textit{id.} at 499.
\item[128.] \textit{Id.} at 605.
\item[129.] Denno, \textit{supra} note 10, at 502.
\item[130.] 536 U.S. 304 (2002).
\end{itemize}
execution of anyone who suffers from mental retardation is barred under the Eighth Amendment’s prohibition of cruel and unusual punishment.\textsuperscript{131}

While neuroscience may offer some answers regarding mitigating circumstances, it is imperative that lawyers utilizing this technology understand how to effectively communicate its findings. For instance, successfully presenting an \textit{Atkins} claim is exceptionally complex.\textsuperscript{132} Not only does “[i]t require[] tremendous preparation involving many hours of consultation with . . . expert forensic psychologists, neuropsychologists, and/or psychiatrists,” it also requires an attorney’s understanding of these concepts and ability to synthesize these issues for presentation to the court and the jury.\textsuperscript{133} The same concerns are true where neuroscience is used to offer evidence of mitigating circumstances.

While such a concern addresses how neuroscience is used, there have also been concerns about who uses neuroscience. Some legal theorists have voiced concern over the possibility of neuroscience being a negative influence in court.\textsuperscript{134} Designating neuroscience as a “double-edged sword,” these commentators have suggested that it “will either get defendants off the hook altogether or unfairly brand them as posing a future danger to society.”\textsuperscript{135} As one commentator put it, “[a] major concern is that prosecutors will seek the death penalty based on neuroscience evidence indicating that a defendant is likely to commit future crimes.”\textsuperscript{136} But in the end, neuroscience has thus far been used only to “provide fact-finders with more complete, reliable, and precise information when determining a defendant’s fate.”\textsuperscript{137}

Lastly, there have been articulated concerns about what neuroscience can show regarding the ingrained tendencies of individuals.\textsuperscript{138} Neuroscience demonstrates that certain brain regions may serve multiple cognitive functions and, vice versa, some cognitive functions may activate different areas of the brain. This knowledge allows neuroscientists to surmise what is going on with the brain and how those cognitive functions affect behavior, all by examining neuroimaging data. Still, just as in any young scientific field, this information is not completely reliable, and it may go so far as to expose information that we were not looking for.\textsuperscript{139} Two specialists in the field, Martha J. Farah and Paul Root Wolpe, stated that

\begin{itemize}
  \item \textsuperscript{131} Id. at 321.
  \item \textsuperscript{133} Id. at 400.
  \item \textsuperscript{134} See, e.g., Denno, supra note 10, at 503.
  \item \textsuperscript{135} Id.
  \item \textsuperscript{136} See Deborah W. Denno, The Place for Neuroscience in Criminal Law, in PHILOSOPHICAL FOUNDATIONS OF LAW & NEUROSCIENCE 69, 74 (Dennis Patterson & Michael S. Pardo eds., 2016).
  \item \textsuperscript{137} See Denno, supra note 10, at 544.
  \item \textsuperscript{138} See Snead, supra note 35, at 1287.
  \item \textsuperscript{139} See id. at 1288.
\end{itemize}
“[a]lthough brainwaves do not lie, neither do they tell the truth.”140 In other words, the fear is that neuroscience could someday “unfairly brand”141 an individual as dangerous. However, our society’s criminal justice system does not punish actions a person may take; it only punishes those that an individual has taken.

Even though there may be gaps in our knowledge of how the brain works, the future holds possibility, and the more we learn about the brain and how it operates, the more the legal field can adapt to and embrace these scientific advancements regarding mitigating circumstances. As for now, commentators and courts have accepted that neuroscience offers a window into the mindset of the individual standing before the court, and the greater a court’s understanding of the individual, the greater the court’s ability to assess and apply an appropriate sentence. This cumulates in greater justice to society and, ultimately, the individual themselves.

II. ANALYSIS

The extensive development of neuroscience and the proliferation of analysis regarding how we do and should use it in a courtroom has captured the legal community’s attention.142 At this point, the question is not whether we should utilize the advancements in technology and insight into defendants that neuroscience offers, the question is when and how.143 However, there is one area of law that has not been thoroughly discussed and could directly benefit from the addition of neuroscience and neuropsychology: sentencing for non-death-penalty-eligible defendants.

To do so, neuroimaging experts contribute to “defendants’ claims that, although legally guilty, they do not deserve to die because the abnormal structure and/or function of their brains diminishes their culpability.”144 As previously discussed, since the 1800s, the criminal justice system has generally embarked on a journey away from mandatory sentences and toward one of discretionary application of sentencing factors, especially in the context of the death penalty.145 Today, public policy is to apply discretionary sentencing based on an individual’s background and

---

140. Id.
141. See Denno, supra note 10, at 494.
142. See Shen, supra note 83, at 357 (“Looking at historical trend[s] in scholarship, it is evident that there has been consistent growth since 2000, strong growth since 2005, and incredibly strong growth in the past two years in the annual number of articles published per year.”); see also id. at 352 (“In the past five years, we have witnessed extraordinary growth in the amount of legal scholarship, legal practice, and public policy at the intersection of law and neuroscience.”).
143. See Denno, supra note 10, at 499; see also Porter v. McCollum, 558 U.S. 30, 40 (2009) (indicating that counsel’s failure to discover and present evidence regarding the defendant’s mental health, mental impairments, family, or military service did not “reflect reasonable professional judgment”). In Porter, a neuropsychologist “concluded that [the defendant] suffered from brain damage that could manifest in impulsive, violent behavior.” Id. at 36.
144. Snead, supra note 35, at 1269.
circumstances, a stance recognized by both Congress and the judiciary. Judges are encouraged to “consider any relevant mitigating evidence regarding the defendant’s character or background, and the circumstances of the particular offense.” Neuroscience can speak to all of these considerations.

A. Neuroscience as a Mitigating Factor at Sentencing

Based on neurological evidence, cognitive neuroscientists are now seeking “to assist defendants’ mitigation claims by invoking cutting-edge brain imaging research on the neurobiological roots of criminal violence” within offenders’ brains. Because neuroscience expands upon how and why a person may act or think the way she does, it can offer insight into a defendant’s true culpability. Whether true culpability should be at the core of sentencing considerations was perhaps best analyzed by Justice Sandra Day O’Connor. In her concurrence in California v. Brown, Justice O’Connor stated, “In my view, evidence about the defendant’s background and character is relevant because of the belief, long held by this society, that defendants who commit criminal acts that are attributable to a disadvantaged background, or to emotional and mental problems, may be less culpable than defendants who have no such excuse.” She questioned whether our criminal justice system should focus more on the individual or focus more on righting a wrong committed against society, asserting that the Supreme Court has been responsive to developing a method for reviewing and applying capital punishment that is “sensible to the uniqueness of the individual.” Although the case she discussed related to capital punishment, she suggested that the Supreme Court’s decisions in Lockett v. Ohio and Eddings v. Oklahoma “reflect the belief that punishment should be directly related to the personal culpability of the criminal defendant.” Justice O’Connor concluded that “the sentence imposed at the penalty stage should reflect a reasoned moral response to the defendant’s background, character, and crime rather than mere sympathy or

146. See Lockett v. Ohio, 438 U.S. 586, 604–05 (1978) (“We recognize that, in noncapital cases, the established practice of individualized sentences rests not on constitutional commands, but on public policy enacted into statutes. The considerations that account for the wide acceptance of individualization of sentences in noncapital cases surely cannot be thought less important in capital cases.”).
149. See, e.g., Stern, supra note 12, at 49 (“Neuropsychological testing plays an important role in explaining to a jury the existence and extent of the plaintiff’s problems with work and other daily activities.”).
151. Id. at 545 (O’Connor, J., concurring).
152. Id. (quoting Eddings v. Oklahoma, 455 U.S. 104, 110 (1982)).
emotion.”156 The Supreme Court has since quoted Justice O’Connor’s sentiments with approval.157

Those sentiments are precisely what lawyers are using to ask courts to consider neuroscience as a mitigating factor at sentencing. Legal minds have argued the following:

[N]eurolaw’s promise to reveal why people think and behave as they do is inescapably built on the idea that people are not agents as the law traditionally views them. Instead, they are guided almost entirely by determined and unconscious chemical cascades which exert irresistible control over an agent’s thinking and behavior.158

Still, the Supreme Court has indicated that while neuroscientific evidence “does not negate” the “responsibility for committing the underlying offense, it does bear upon an argument” that a mental illness or problem may impair the “ability to conform . . . conduct to the requirements of the law.”159 Thus, one strategy that defense attorneys may choose to utilize is to affirmatively demonstrate through neuroscience that, while an individual may be guilty, she should be held less culpable for her actions than other members of our society would.160 Indeed, the Court has long recognized that individualized sentencing is appropriate, even though it is not constitutionally required.161 Moreover, a judge’s expansive discretion at sentencing affords her the leeway necessary to consider vast amounts of information,162 including information about a defendant’s background and life choices.163

Neuroscience can provide a qualified assessment of how culpable society may want to hold a particular person, given their background and its effect on their abilities to process situations in accordance with societal norms.164 Although the goal of the Sentencing Guidelines is to ensure more uniform sentences, neuroscientific evidence fits into the mold created by the Sentencing Guidelines. Even if a judge chooses not to depart from the Guidelines range for a given defendant, she may consider the weight of the

156. Id.
158. Steven K. Erickson, Blaming the Brain, 11 MINN. J.L. SCI. & TECH. 27, 57 (2010).
159. Bell v. Thompson, 545 U.S. 794, 810 (2005); see also Lockett v. Ohio, 438 U.S. 586, 602 (1978) (“[T]he definition of crimes generally has not been thought automatically to dictate what should be the proper penalty.”).
161. See Lockett, 438 U.S. at 602.
162. Id. (“[S]entencing judges traditionally have taken a wide range of factors into account.”).
163. Id. at 602–03 (“And where sentencing discretion is granted, it generally has been agreed that the sentencing judge’s ‘possession of the fullest information possible concerning the defendant’s life and characteristics’ is ‘[h]ighly relevant—if not essential—to the selection of an appropriate sentence . . . .’” (alteration in original) (emphasis omitted) (quoting Williams v. New York, 337 U.S. 241, 247 (1949)).
164. See Porter v. McCollum, 558 U.S. 30, 41 (2009) (“This is not a case in which the new evidence ‘would barely have altered the sentencing profile presented to the sentencing judge.’ The judge and jury at Porter’s original sentencing heard almost nothing that would humanize Porter or allow them to accurately gauge his moral culpability.” (citation omitted) (quoting Strickland v. Washington, 466 U.S. 668, 700 (1984)).
evidence in applying a sentence from the low or the high end of a given Guidelines sentencing range. In short, the ability of neuroscience to explicitly identify and explain how one’s background affects culpability is supplemental information a judge may find useful at sentencing in attempting to distribute justice.

B. Combating Implicit Bias with Neuroscience

Neuroscience can also be used to combat any implicit bias that may be present at sentencing. Implicit bias, or social cognition, is the process by which the brain uses “mental associations that are so well-established as to operate without awareness, or without intention, or without control.”165 There has been an explosion of research over the past decade about implicit bias and its relationship to decisions and actions. This research offers new avenues and opportunities to intervene and reduce the effects of bias in institutions and interactions. Evidence from hundreds of thousands of individuals shows the following:

1. The magnitude of implicit bias toward members of outgroups or disadvantaged groups is large,
2. implicit bias often conflicts with conscious attitudes, endorsed beliefs, and intentional behavior,
3. implicit bias influences evaluations of and behavior toward those who are the subject of the bias, and
4. self, situational, or broader cultural interventions can correct systematic and consensually shared implicit bias.166

There are three important aspects regarding implicit biases that are crucial to remember. First, we all have biases; they are a way for us to process and organize the vast amounts of information that we observe every day.167 Second, unconscious biases often conflict with society’s egalitarian values.168 Third, implicit biases often predict and determine actions and decisions more so than the values that we make sure to explicitly adopt.169 Thus, implicit biases are always present, including at sentencing.

Researchers have analyzed the effect of a trial court judge’s implicit biases on defendants in the judge’s courtroom. One study found that, according to the Implicit Association Test, consistent with other Americans, judges held implicit associations concerning African Americans.170

165. FAQs, PROJECT IMPLICIT, https://implicit.harvard.edu/implicit/demo/background/faqs.html (last visited Oct. 16, 2016) [https://perma.cc/YSR3-BER7].
168. See id. at 1222.
169. See id.
170. See id.; see also Jerry Kang et al., Implicit Bias in the Courtroom, 59 UCLA L. REV. 1124, 1143 (2012) (“Put another way, data show that when the race of the defendant is explicitly identified to judges in the context of a psychology study . . . judges are strongly motivated to be fair, which prompts a different response from White judges (who may think to themselves ‘whatever else, make sure not to treat the Black defendants worse’) than Black judges (who may think ‘give the benefit of the doubt to Black defendants’). However, when
However, these associations were only influential when the race of the defendant was manipulated through subliminal techniques. When the race of the defendant was explicitly identified, implicit associations had no influence on judgment. These results suggest that judges are able to control the influence of unconscious racial bias but only when they are focused on doing so.

Because implicit biases play a role in the courtroom and at sentencing, the question then becomes what should we do about it? Neuroscience offers insight into practical ways to reduce bias by offering mitigating evidence that may help the sentencing judge identify these biases. When a neurological test can demonstrate that a defendant’s thought process does not operate as one a judge would normally encounter—whether that is due to diminished culpability or a brain injury—a judge could consider that information, counteracting any implicit biases she may be holding about the defendant’s obvious characteristics. After all, reliance on physical evidence that can be brought forth to affirmatively show that someone thinks differently is immensely more convincing than the mere acknowledgment that an individual experienced childhood trauma or had a brain injury.

Neuroscience also contributes to one of the main ways to combat implicit biases: individuation. Individuation requires the person attempting to combat her implicit biases to gather specific information about an individual before her. Thus, instead of making judgments on the basis of the defendant’s group characteristics, a judge armed with neurological information and data can make judgments based on the defendant’s personal characteristics. In other words, the judge would be able to explicitly recognize that a group characteristic that may result in a bias is merely one of that individual’s many, many attributes.

Individuation walks hand in hand with another common way to combat implicit bias: perspective-taking. Perspective-taking constitutes “imagining oneself in the shoes of someone from a different social or ethnic

race is not explicitly identified but implicitly primed... perhaps the judges’ motivation to be accurate and fair is not on full alert.”

171. See Rachlinski, supra note 167, at 1223.
172. Id.
173. See generally Kang et al., supra note 170.
174. Kang & Banaji, supra note 166, at 1064 (indicating that implicit biases can be counteracted through self, situational, and broader cultural interventions).
175. See Snead, supra note 35, at 1313 (“It is one thing to deny that human decision-making is purely mechanical when your opponent offers only a general, philosophical argument. It is quite another to hold your ground when your opponent can make detailed predictions about how these mechanical processes work, complete with images of the brain structures involved and equations that describe their function.” (quoting Joshua Greene & Jonathan Cohen, supra note 18, at 1781)).
177. See id. at 1270.
178. Id.
179. Id.
When individuals take part in perspective-taking viewpoints, studies show that it “weakens the automatic expression of racial biases.” In fact, various “perspective-taking activities substantially decrease[] implicit bias as measured by the IAT and behavioral changes.” Neuroscience, and especially its subfield of study, neuropsychology, with its in-depth reviews of a person’s background and past experiences, can help a judge engage in perspective-taking. Further, it exposes judges to a defendant’s background, as discussed by a relevant expert, as opposed to the defendant herself or the defendant’s attorney. Establishing a connection with the defendant’s background and attempting to further understand it “may help a judge take, rather than evade, responsibility for the consequences of her decisions.”

There is no shortage of ways that neuroscience can facilitate the sentencing process. As the science continues to develop, criminal justice system actors should continue to look for ways in which it can provide even more insight and clarity into a defendant’s life, providing a fuller picture for sentencing purposes.

CONCLUSION

Since the seventeenth century, the criminal justice system has concerned itself with mental states of the accused. Even though the concept of neuroscience has only been around since the 1960s, recent expansions and developments in the field have opened a window, shedding light on the vast darkness that encompasses what little we know about the brain. It indicates that an offender’s criminal intent or mental state may be the product of her past experiences, especially those during childhood. Using this vital information during the sentencing phase of a convicted offender could more thoroughly help a judge apply an appropriate sentence and avoid biases, thereby providing more principled justice.

Courts have the tools necessary to handle the newest forms of technology, even those in their infancy. It will be up to the lawyers appearing before the court to learn and utilize the many offerings of neuroscience in both criminal and civil trials and at sentencing hearings as a form of mitigating evidence. It will be up to judges to contemplate the information offered to them and employ and utilize it appropriately, both in sentencing and to counteract their own implicit biases. And it will be up to state and federal legislators to take advantage of the information and data discovered by these studies and form more thoughtful and responsive

181. Id.
182. Id.
183. Id. at 738.
185. Id.
186. See id. at 505.
policies and statutes that address the systemic problems leading to mass incarceration.

Defendants exposed to childhood trauma are one of the most prominent examples of those who can benefit from utilizing neuroscientific evidence in a courtroom. Individualized assessments of a defendant’s background can expand on why she may have behaved a particular way, such as if it was the result of repeated exposure to traumatic situations during childhood. This particularly applies to, for example, those who grow up in dangerous inner-city neighborhoods and those who end up in the foster care and welfare systems. Not only does neuroscience offer judges insight into individuals such as these, but it can also facilitate judges’ attempts to counteract implicit biases.

In sum, participants in the criminal justice system—from judges and defense attorneys to prosecutors and legislators—should keep abreast of the developments in neuroscience and consider whether the use of neuroscience and neuropsychology would be beneficial given the particular circumstances of the case, either in dealing with a particular individual or a societal issue. As the technology continues to flourish, we should embrace it in our communal efforts to continue bending the arc of the moral universe toward justice.187

Neurocriminology: implications for the punishment, prediction and prevention of criminal behaviour

Andrea L. Glenn and Adrian Raine

Abstract | Criminal behaviour and violence are increasingly viewed as worldwide public health problems. A growing body of knowledge shows that criminal behaviour has a neurobiological basis, and this has intensified judicial interest in the potential application of neuroscience to criminal law. It also gives rise to important questions. What are the implications of such application for predicting future criminal behaviour and protecting society? Can it be used to prevent violence? And what are the implications for the way offenders are punished?

Advances in neuroscience are increasing our understanding of how our biology influences our behaviour — for both good and bad. The emerging field of neurocriminology seeks to apply techniques and principles from neuroscience to improve our understanding of crime, to predict crime and ultimately to prevent crime. Such an approach would have the potential economic and social benefits that are associated with violence reduction, but it also raises neuroethical concerns.

In this Perspective article, we discuss the current state of research in neurocriminology. We provide an overview of the neurobiological abnormalities that are associated with criminal behaviour and consider the genetic and environmental factors that may contribute to these abnormalities. We highlight studies conducted to date, many of which suggest that biological factors may aid in the prediction of future crime and violence. We then discuss implications of this research in the legal system.

The current state of neurocriminology

There are now relatively extensive literatures that document relationships between antisocial behaviour and biological functioning. With some exceptions, most studies are correlational and cross-sectional, and largely do not provide information on specific genetic or environmental factors that may mediate these relationships. However, an increasing number of prospective longitudinal studies are examining whether the presence of specific biological factors, such as hormone levels, neurotransmitter levels, physiological indices or brain impairments, is predictive of future offending. Because most studies define antisocial behaviour and crime broadly, without distinguishing between violent and non-violent offenders, this article largely concerns the broad propensity to criminal behaviour.

Genetics. Results from well over 100 behavioural genetics studies with different designs — including twin studies, studies of twins reared apart and adoption studies — have converged on the conclusion that antisocial and aggressive behaviour have a considerable genetic basis. Estimates of the variance that is attributable to genetics vary, but several meta-analyses place the level at between 40–60%. Heritable influences, with some exceptions, are broadly consistent across gender and ethnicity. Adoption studies in particular have the advantage of being able to truly separate genetic from environmental factors and provide a converging line of evidence that there are heritable influences on antisocial and aggressive behaviour. A recent systematic review of over 100 studies showed that the proportion of variance in antisocial and aggressive behaviour explained by genetic factors is variable.

Recently, research has focused on identifying which specific genes confer risk of antisocial behaviour. Several genetic variants that incrementally increase the risk of antisocial behaviour have been identified. Although approximately one-half of 185 studies have reported positive findings, a meta-analysis revealed that no variant was associated with aggression at the 5% level of significance. This finding underscores the idea that, as with other complex behaviours, the contribution of any single gene to antisocial and aggressive behaviour is likely to be quite small. It is possible that a combination of a larger number of gene variants substantially increases the risk of aggressive behaviour. Nevertheless, knowledge about individual genes may prove to be useful in improving our understanding of the mechanisms and pathways that increase the risk of antisocial behaviour. Importantly, the environment plays an equally influential part; indeed, some genetic variants confer risk of antisocial behaviour only in the presence of particular environmental risk factors, such as abuse in early childhood. Research in epigenetics has shown that the environment can influence how genes are functionally expressed in an individual (and even in specific brain areas); this finding undermines traditional arguments of biological determinism.

Prenatal and perinatal influences. Early health risk factors, sometimes in conjunction with social risk factors, have been found to be associated with an increased probability that a young infant will develop antisocial and aggressive behaviour. During the prenatal and perinatal period, a number of factors may be important. Birth complications, in combination with maternal rejection of the child in the first year of life, have been associated with violent criminal offending at the age of 34 years in a study carried out in Denmark. This predictive finding has been replicated in the United States, Canada, Sweden and Finland with respect to violence in adulthood, and in Hawaii and Pittsburgh (USA) with respect to childhood antisocial behaviour. Five other studies have shown associations between birth complications and externalizing-behaviour problems (such as aggression, delinquency and hyperactivity) in children. Fetal maldevelopment during the second trimester of pregnancy, as indicated by minor physical anomalies in the child (such as low seated ears or a single palmar crease), has been associated with later violent delinquency and violent offending in adulthood. The association between fetal neural maldevelopment and childhood aggression and adolescent conduct disorder may be even more pronounced when combined with effects of poor parenting or social adversity. Criminal offending and psychopathy have been associated with another indicator of disruption in fetal development — namely, cavum septum pellucidum, which in about 60% of cases is absent. This finding underscores the idea that, as with other complex behaviours, the contribution of any single gene to antisocial and aggressive behaviour is likely to be quite small. It is possible that a combination of a larger number of gene variants substantially increases the risk of aggressive behaviour. Nevertheless, knowledge about individual genes may prove to be useful in improving our understanding of the mechanisms and pathways that increase the risk of antisocial behaviour. Importantly, the environment plays an equally influential part; indeed, some genetic variants confer risk of antisocial behaviour only in the presence of particular environmental risk factors, such as abuse in early childhood. Research in epigenetics has shown that the environment can influence how genes are functionally expressed in an individual (and even in specific brain areas); this finding undermines traditional arguments of biological determinism.
Maternal nicotine consumption and alcohol consumption during pregnancy are also factors that may predispose individuals to violent offending in adulthood — findings that have been replicated across many studies in several continents. Even small amounts of alcohol during pregnancy (one drink per week) have been associated with increased childhood aggression in the offspring. There is current debate regarding whether nicotine exposure predisposes to offending by causing fetal hypoxia that results in brain impairment or whether this association is genetically mediated.

Lead levels have been associated with juvenile delinquency and aggressive behaviour in at least six studies. From a prospective viewpoint, high lead levels in the mother during the first and second trimester of pregnancy are associated with an increased risk of being arrested for violent crimes in adulthood. High dentine lead levels assessed at the ages of 6–9 years have been associated with increased violent offending at the ages of 14–21 years, and poorer cognitive functioning mediates this relationship. Some studies that carefully controlled for potential confounds such as poverty, maternal smoking, alcohol use and drug use have shown that these findings apply to women as well as men. Higher manganese levels in the mother during pregnancy have also been associated with increased externalizing-behaviour problems (defined as aggressive, destructive and defiant behaviour) in children aged 8–9 years.

Poor nutrition in either the first or second trimester of pregnancy has been associated with a 2.5-fold increase in antisocial personality disorder in the offspring. Malnutrition in infancy is associated with conduct problems in adolescence, a relationship that is partly mediated by low IQ. Similarly, children with signs of malnutrition at the age of 3 years have much higher rates of aggressive and antisocial behaviour at the ages of 8, 11 and 17 years over and above any contribution from social risk factors. This relationship is also mediated by low IQ.

Together, these findings suggest that a number of early environmental factors may increase the risk of antisocial behaviour as late as adulthood, probably via effects on biological systems. Hormones and neurotransmitters. The steroid hormones cortisol and testosterone have been the most intensively researched hormones in relation to antisocial behaviour. Disruptions in the hypothalamic–pituitary–adrenal (HPA) axis, the body’s stress response system that regulates the release of the hormone cortisol, are frequently observed in antisocial people. Associations between antisocial behaviour and cortisol levels vary depending on the type of antisocial behaviour and other factors. Psychological stress at various stages during development may produce lasting changes in HPA axis functioning and thereby predispose an individual to antisocial behaviour. Low levels of cortisol in childhood are predictive of aggressive behaviour 5 years later, in adolescence. Similarly, a study showed that boys who were identified as having behavioural problems and who had low cortisol levels showed more aggressive behaviour at a follow-up assessment 2 years later.

Increased testosterone levels have been repeatedly associated with increased aggressive behaviour in adults. Caveats include the fact that this relationship appears to be less evident in pre-pubertal individuals, and meta-analyses of this relationship find a small effect size of R = 0.08 (Ref. 54). Some randomized, placebo-controlled crossover trials have shown that testosterone administration increases aggressive behaviour in adult males, which is suggestive of a causal connection, although other experimental studies using lower doses of testosterone did not show an increase in aggressive behaviour. Increased levels of testosterone at the ages of 10–12 years are predictive of assaultive behaviour at the ages of 12–14 years, norm-violating behaviour at the age of 16 years and cannabis use at the age of 19 years. Higher levels of testosterone at the age of 16 years are associated with crime in adulthood. Multiple neurotransmitter systems have been implicated in aggression, and the best-replicated correlate of human aggressive behaviour is a low level of serotonin. Low levels of serotonin in cerebrospinal fluid are a particular marker of people who show impulsive aggressive behaviour. An experimental manipulation that reduces serotonin levels in the brain (that is, acute tryptophan depletion) reduced functioning of the orbitofrontal cortex during an inhibitory motor control task, a region commonly implicated in antisocial behaviour. However, aggression has also been associated with reduced monoamine oxidase A (MAOA) levels in the brain. MAOA is an enzyme that breaks down serotonin and other neurotransmitters, and hence lower levels of MAOA would presumably result in higher serotonin levels. This seemingly contradictory finding demonstrates the need for studies that simultaneously examine multiple biological markers in order to obtain information about how neurotransmitters may interact with each other to increase the risk of aggression.

Psychophysiology. Psychophysiological differences have also been observed between antisocial groups and control groups. Meta-analyses and reviews conclude that low resting heart rate is probably the best-replicated biological correlate of antisocial and aggressive behaviour in children and adolescents. Low resting heart rate may indicate a lack of fear and a reduced likelihood of experiencing negative affect in response to a criminal act. Low heart rate in childhood and adolescence has been associated with adult crime in all four longitudinal studies conducted to
date\(^{46}\). Across these studies, low resting heart rate was found to be as strong a predictor of future antisocial behaviour as it is for current antisocial behaviour\(^{45}\). In delinquents who were arrested for a minor offence at the age of 14 years, attenuated heart rate responses to a stressor were associated with both a shorter time to re-offend as well as with a greater number of re-offences within a 5-year period\(^{45}\). Another study showed that, after multiple confounds had been controlled for, low heart rate at the age of 18 years predicted higher conviction frequency and higher levels of violence up to the age of 50 years\(^{46}\).

Psychophysiological indicators of under-arousal — such as slow-frequency electroencephalographic activity and reduced skin conductance activity — at the age of 15 years are predictive of crime at the age of 24 years\(^{47}\). A recent meta-analysis\(^{48}\) has documented a reduced amplitude of the P300 event-related brain potential, which is thought to reflect inefficient recruitment of neural resources during information processing, in adult antisocial populations. Similarly, a reduced P300 amplitude at the age of 11 years has been associated with criminal offending at the age of 23 years. P300 amplitude predicted offending at the age of 23 years over and above measures of antisocial behaviour at the age of 11 years\(^{51}\).

Poor autonomic fear conditioning — the ability to learn associations between neutral cues and aversive stimuli — is another well-replicated correlate of adult criminal and psychopathic adult offending\(^{52,53}\), conduct disorder in children and adolescents\(^{54,55}\), and juvenile offending\(^{56}\). A review of 46 human brain imaging studies suggests that deficits in fear conditioning may reflect abnormalities in a common core fear network that consists of the amygdala, insula and anterior cingulate\(^{57}\). Indeed, numerous brain imaging studies find abnormalities in these areas in antisocial people, although this has been disputed with respect to individuals with psychopathic traits\(^{58}\) — a specific subgroup of criminal offenders. Poor electrodermal fear conditioning at the age of 3 years is associated with convictions for criminal offences at the age of 23 years\(^{45}\). In addition to aiding in the prediction of future offending, individual differences in fear conditioning may also provide information about which antisocial individuals may desist from future violence or be particularly amenable to treatment. For example, adolescents who were identified as being likely to commit crimes in adulthood by virtue of being antisocial at the age of 15 years but who did not go on to develop into adult criminal offenders at the age of 29 years showed superior fear conditioning compared with both antisocial adolescents who become offenders and non-criminal controls\(^{59}\).

**Brain imaging and neurology.** Reduced functioning in the frontal lobe of the brain is to date the best-replicated brain imaging correlate of antisocial and violent behaviour. In particular, a meta-analysis of 43 structural and functional imaging studies found that the largest reductions in structure and function within the frontal lobe of antisocial individuals were observed in the orbitofrontal cortex, anterior cingulate cortex and dorsolateral prefrontal cortex\(^{61}\). The dorsolateral prefrontal cortex is associated with self-regulatory processes, including attention and cognitive flexibility, and may be linked to the antisocial features of impulsivity and poor behavioural control\(^{61}\). The anterior cingulate is involved in error processing, conflict monitoring and avoidance learning\(^{62-64}\). Individuals with damage to this region are more disinhibited and aggressive\(^{65}\), and demonstrate impairments in inhibitory control and emotion processing\(^{66,67}\). The ventral prefrontal cortex, including the orbitofrontal cortex, has received particular attention given its role in emotion processing, learning from reward and punishment, and decision making\(^{68,69}\).

The possibility of a causal connection between impaired orbitofrontal cortex structure and/or function on the one hand, and crime and/or violence on the other, has been raised by neurological studies showing that head injury in ostensibly normal individuals precedes the onset of disinhibited antisocial behaviour. For example, higher levels of aggression were found in war veterans who had experienced penetrating head injuries that were localized to the ventral prefrontal cortex\(^{70}\). Furthermore, neurological patients who had suffered from an accidental head injury to the ventral prefrontal cortex show poor decision making, reduced autonomic reactivity to socially meaningful stimuli and psychopathic-like behaviour\(^{44}\). In a particularly striking example, a tumour in the orbitofrontal region preceded the onset of paedophilia in an individual; after resection of the tumour, the person’s behaviour returned to normal\(^{71}\) (BOX 2).

The amygdala is another brain region that is consistently identified as showing altered activity in brain imaging studies of antisocial individuals. The type of deficit may vary in different subgroups of antisocial individuals. Adults and youths with psychopathic traits, who have blunted emotional responding and may engage in more cold, calculated aggression, have reduced amygdala volume\(^{72}\) and functioning\(^{73-77}\), whereas individuals with a more impulsive, reactive form of aggression demonstrate exaggerated amygdala reactivity\(^{78}\). Reduced amygdala volume in psychopathic individuals has been localized to the basolateral, lateral, cortical and central nuclei — regions that are involved in emotion processing, fear conditioning and autonomic reactivity to affective stimuli\(^{72}\). Of note, patients with damage to the amygdala have a reduced sense of danger, are less fearful\(^{77}\) and have deficits in the recognition of fearful facial expressions\(^{79}\) (a process involved in experiencing empathy). The association noted earlier\(^{80}\) between poor classical conditioning in childhood and crime in adulthood suggests, but does not prove, a causal relationship between amygdala functioning and antisocial behaviour.

Most brain imaging studies are essentially correlational and cross-sectional, and until recently no longitudinal brain imaging research on antisocial populations has been conducted. Two recent studies have indicated the potential for neuroimaging to provide incremental predictive power in predicting re-offending. One study showed that reduced functioning in the anterior cingulate during a go/no-go task in prisoners doubled the likelihood of re-arrest 3 years later\(^{79}\). A second study of high-risk community males showed that reduced amygdala volume at the age of 26 years was associated with violent offending 3 years later\(^{81}\). As has been observed in other biological longitudinal research, both studies showed predictive utility of brain measures over and above past history of antisocial behaviour and other confounds.

Other longitudinal studies have shown that incurring brain damage increases the risk of criminal behaviour. A longitudinal study of 231,129 individuals from Sweden documented a threefold increase in violent crime after traumatic brain injury (TBI) after adjusting for demographic confounds\(^{81}\). A prospective longitudinal study of 12,058 individuals from Finland showed that TBI during childhood and adolescence was associated with a 1.6-fold increase in crime in adulthood after controlling for confounds; children suffering from TBI before the age of 12 years started their criminal careers significantly earlier than those who suffered from TBI after the age of 12 years\(^{81}\). These studies demonstrate that information about brain structure and function, regardless of whether the origins are neurodevelopmental or a result of a direct physical insult later in life, may be of some use in identifying which individuals are at an increased risk of criminal behaviour.
Cross-sectional brain imaging studies are correlational and cannot prove a causal association. Individual case studies can, however, be suggestive of causality.

Michael was a 40-year-old schoolteacher and past correctional officer. He was happily married to his wife and loved both her and Christine, his stepdaughter. He had no prior history of criminal or deviant behaviour. However, Michael began to change. He became uncharacteristically aggressive with his wife and began taking pornography to school. His bedtime rituals with his pre-pubescent stepdaughter, which had previously consisted of singing lullabies, became more sordid, and he eventually got into bed with her. He was found out and convicted of child molestation.

Michael had to decide between a prison sentence and a treatment programme. He chose the treatment programme but was expelled after propositioning female staff. The night before he was due to be transported to prison, he went to the emergency room complaining of a severe headache. There he continued to solicit sexual favours from staff.

An astute neurologist ordered an MRI scan after Michael wet his trousers without showing any apparent concern. The MRI revealed a tumour growing from the base of the orbitofrontal cortex (see the figure, which shows MRI scans of Michael’s brain at the time of the initial neurological evaluation, revealing a tumour mass displacing the right orbitofrontal cortex). After the tumour was resected, Michael’s behaviour returned to normal, and he was reunited with his wife and stepdaughter. After several months of normal behaviour, his wife discovered child pornography on his computer. Michael was re-examined, and it was discovered that the tumour had regrown. It was resected for a second time, and for at least 6 years after the resection Michael’s behaviour has returned to normal.23

The case comes almost as close as one can get to a causal connection between ventral prefrontal brain pathology and deviant behaviour — a pendulum moving from normality to brain dysfunction to paedophilia to neurosurgery to normality, and back again. In the face of the order in which events occurred, was Michael responsible for his inappropriate sexual behaviour with his stepdaughter?

Figure is reproduced, with permission, from REF. 71 © (2003) American Medical Association.

Remaining challenges for research on biological risk factors for violence and crime. In sum, in recent years, evidence of the importance of biological factors in antisocial behaviour has accumulated and is being recognized as valuable in our understanding of crime and violence. With advances in neuroscience and the design of longitudinal investigations, studies are becoming methodologically stronger. Taken together, it is becoming increasingly harder to argue that biological factors do not predispose some individuals to adult crime. This conclusion neither diminishes nor replaces social and environmental perspectives on crime causation.14,29,40 Together, genetic and environmental factors shape the way that biological systems develop and function, and thus affect multiple complex psychological processes that are important in controlling and regulating behaviour and in behaving morally.

Important gaps in our knowledge remain. Very little is known about the neurobiology of regulatory crimes, and one study showed increased cortical thickness and better prefrontal functioning, as opposed to impairments, in white-collar criminals.85 A future challenge in neurocriminology lies in parsing out the specific genetic and environmental influences that induce neurophysiological changes that result in the more proximal cognitive, affective and behavioural risk factors for violence. In other words, what neurobiological processes mediate the relationship between the well-documented early social risk factors and violence in adulthood?

A few studies have begun to explore how genetic and environmental factors affect the brain. For example, researchers have found that the adolescent offspring of mothers who smoke during pregnancy have reduced thickness in two regions of the brain that have been implicated in antisocial behaviour — the orbitofrontal cortex and middle frontal cortex.83 Children exposed to high levels of lead early in life have been shown in adulthood to have reduced grey matter volume in the brain, particularly in the prefrontal cortex.84 Males with a common polymorphism in the MAOA gene (which is present in about 30% of the population) have an 8% reduction in the volumes of the amygdala, anterior cingulate and orbitofrontal cortex,85 which suggests that there is a causal pathway from genes to brain to antisocial behaviour.

A predisposition to criminal behaviour is unlikely to be reduced to one or even two simple brain circuits but probably involves multiple brain dysfunctions and multiple circuits that each give rise to different risk factors for violence. Thus, the future use of brain imaging in the assessment of risk of criminal behaviour will require a much more sophisticated understanding of these circuits. Although brain imaging techniques have advanced rapidly in the past few decades, there are still many limitations to these methods.86 However, with continued methodological improvements in neuroscience research, we will gain more information about how brain regions function together to predispose individuals to criminal behaviour.

Although only a few prospective studies have been conducted, findings from research on early risk factors suggest that information about biological factors in youths may aid in the prediction of which individuals are more likely to engage in crime and violence later in life. Such information may also help to identify individuals who are particularly amenable to rehabilitation. In a review of ten studies implementing variants of cognitive–behavioural therapy in individuals with antisocial behaviour,87 multiple neurobiological factors were predictive of treatment response and progress, including heart rate, hormone levels and neuropsychological measures of risk taking, sensitivity to negative consequences, impulsivity, cognitive flexibility and emotion processing. Although such initial findings are provisional, these neurobiological characteristics could ultimately help to determine which offenders are best suited to specific rehabilitation programmes and are more likely to re-integrate into society.
safely. A major challenge that remains to be addressed is the identification of socially acceptable psychosocial or biological intervention programmes that target biological risk factors for criminal behaviour.

The legal context
Neurocriminology interfaces with the judicial system at three main levels: punishment, prediction and prevention. To what extent does the growing body of knowledge on the neuroscience of crime and violence suggest that we should rethink our approach in these three domains? Although it is unlikely that neurocriminology will result in any radical or swift shift in the operation of the criminal justice system in the very near future, it is not inconceivable that some modest change may occur in these areas at some point, assuming that the field continues to develop and evolve, as the past two decades have suggested.

Punishment. Punishment is predicated on blameworthiness, and the extent to which we blame individuals is a function of the extent to which they can be held accountable for their actions. Such accountability in a legal context is based on the concept of responsibility.

In this context, let us assume that, to some extent, neurobiological abnormalities or insults relatively early in life predispose some individuals to a life of crime and violence. We also assume that offenders are not responsible for being exposed to these early risk factors for violence. So are these offenders responsible for their behaviour, and if so, to what degree? In the case of Michael, in whom a tumour in the orbitofrontal region preceded the onset of paedophilia, which disappeared after the tumour had been removed (BOX 2), the question was: was he responsible for his paedophilia? Currently, in the United States, an individual is deemed ‘responsible’ for their actions if two conditions are met: first, they have sufficient rational capacity; and second, they are not acting under coercion. Rational capacity is typically interpreted as whether the individual knew what he or she was doing and understood that his or her actions would have consequences. Michael’s (BOX 2) is a telling case because the temporal ordering of events — from normality to brain tumour to paedophilic interest to tumour resection to normality, and back again — is suggestive of causality in this particular case. However, in his own words, Michael admitted, ‘...somewhere deep, deep, deep in the back of my head, there was a little voice saying “You shouldn’t do this”‘ (REF. 88). He knew at the time of the act what he was doing, and he also knew that what he was doing was wrong. In the eyes of the law, Michael was legally responsible for his actions.

Given that Michael would be considered legally responsible, it is difficult to argue that someone with a less obvious neurobiological ‘predisposition’ to offending than that of Michael — such as reduced functioning of the amygdala during a moral decision-making task, carrying a specific variant of the MAOA gene or a significant but non-obvious volume reduction in prefrontal grey matter — is not responsible for his or her actions. In most criminal cases, the causal flow from biological risk to offending will never be known. Behaviour has a cause, and identifying the neural basis of a behaviour in an individual does not in itself establish that the individual had diminished rational capacity91. Therefore, as the law currently stands in the United States and other countries, the documentation of neurobiological risk factors, no matter how early they originated, does not render that individual lacking responsibility.

Despite this current legal stance, a challenging question concerns whether the current law pertaining to responsibility is in need of modest revision. This is ultimately a normative question over which there can be reasonable disagreement. Even without invoking the presence of biological risk factors to suggest impaired rational capacity, it has been argued that severe psychopaths should not be held responsible on the grounds that they have no sense of moral rationality — they are not sensitive to moral concerns and thus do not have the same moral sense as most people in society.90. When one considers in addition an increasing body of evidence showing that neurobiological factors contribute to criminal psychopathy in adults as well as to behaviour in children with psychopathic-like traits, such a revision perhaps becomes more compelling, particularly in a case in which an individual has several documented neurobiological and psychosocial risk factors for violence potential, as in the case of Donta Page (BOX 3).

The judicial system acting in a practical world essentially conducts binary decision making: for example, in establishing innocence versus guilt. Determination of diminished capacity in the United States similarly involves a categorical judgement on the presence or absence of a mental disability. The fairness of this binary judgement can be reasonably questioned. The widespread consensus of experts is that crime and antisocial behaviour are dimensional, and not categorical, constructs.92. Risk factors associated with antisocial and criminal violence are also usually dimensional in nature (for example, the degree of prefrontal dysfunction and the resting heart rate), although some may be categorical (for example, the presence of TBI or genetic polymorphisms). Unlike in the United States, the judicial practice in the Netherlands is guided by a five-point scale for assessing the degree of criminal responsibility, with evaluations including personality and neuropsychological testing93. Thus, although neuroscience has no current definitional bearing on concepts of responsibility, it is not without international precedent to consider a revision to legal practice in the United States, United Kingdom and other countries so that responsibility may, in the future, be assessed on a continuum using measures that include neurobiological variables.

Although a sensible dividing line needs to be drawn for practical reasons, in theory one can conceive of a set of multiple neurobiological and genetic influences that, combined with social influences, diminish responsibility to varying degrees. To the extent that neuroscience provides reliable methods to document these influences objectively, and assuming that methodologies become less expensive and quicker and easier to implement than hitherto, we anticipate that responsibility will eventually be conceptualized more broadly than it is today. For example, although cognitive intelligence is the benchmark used by the law to document the capacity for rationality, the relatively new fields of affective psychology and affective neuroscience are providing us with evidence that emotion informs decision making — a finding that is not yet instantiated in the law. Can individuals therefore be fully responsible when the feeling for what is moral is diminished? What may be just as important as knowing the difference between right and wrong when making moral decisions is having the feeling of what is right and wrong. As recent studies have documented in psychopaths, some individuals may have deficits in brain regions that are important for generating these emotional responses.

The facts that research in the field uniformly recognizes substantial affective impairments as a core feature of psychopathy and that there is no longer any reasonable doubt that such affective impairment influences behaviour95,96 raise the question of whether the legal system will eventually reformulate its current, long-standing concept of responsibility. For example, environmental head injuries can change an otherwise responsible individual into a person...
who, although cognitively capable of differentiating right from wrong, lacks the neural regulatory affective and behavioural control over their behaviour. It has been suggested that as neuroscience begins to offer a more detailed and specific account of the physical processes that can lead to irresponsible or criminal behaviour, the public perception of responsibility may begin to change in the same way that public viewpoints on addiction have shifted from addiction as a failure of personal responsibility towards addiction as a disease.

**Prediction.** If biological factors could predict future violence over and above predictions based on social variables, even opponents of a neuroscientific perspective on crime would have to agree that neurobiology has added value in this area. Whether or not such biological factors are causes or merely correlates of violence is irrelevant to the issue of prediction — the fact that they add predictive value is the currency of risk assessment in prisoners who are about to be released.

Given that approximately 50% of the variance in aggressive and antisocial behaviour can be explained by genetic influences, a compelling case could be made for using biological information to improve violence prediction. However, the fact that molecular genetic studies have so far largely failed to identify specific genes that can account for more than 1% of the variance in any complex behavioural trait gives considerable pause for thought. Molecular genetic advances have, in theory, the potential to elucidate and identify specific genetic factors that predispose individuals to crime in the future, but currently the value of genotyping individuals to predict future violence is limited.

Perhaps surprisingly, endophenotypes such as prefrontal dysfunction and low heart rate, which reflect compound genetic and environmental influences, may currently explain more of the variance in adult violence than any individual genotype and may have more traction in predicting future violence. The literature reviewed above has revealed several replicable early biological correlates of later violence. Some studies have shown that neurobiological markers can predict, over and above well-replicated psychosocial risk factors, which individuals will demonstrate antisocial or psychopathic traits. The two recent imaging studies described above, together with multiple studies that have identified psychophysiological and hormone predictors of future offending, provide some support for the conclusions made in a Royal Society report that neuroscience may have future value in predicting re-offending.

Despite the potential promise, and indeed likelihood, that neurobiology could provide at least modest increases in predictive power, methods used to predict the potential of future re-offending in about-to-be-released prisoners have so far never incorporated neurobiological markers into the risk assessment equation. There are three main reasons for this. First, the evolving body of knowledge on neurocriminology has not yet been accepted in the social sciences and among practitioners. Second, neurobiological measures are less easy to collect than behavioural, social and psychological data. Third, there have been long-standing ethical concerns about collecting biological data on offenders. This may change given that DNA is now collected on all arrestees in the United States. Technical developments are also increasingly making neurobiological risk assessments more feasible and practical, and some, such as the measurement of resting heart rate, are already incorporated into standard medical practice at the community level.

Any major advances in predicting future violence will be based not just on progress in neurocriminology but also on statistical advances. Machine-learning techniques such as random forest have already been documented to improve the prediction of future charges of homicide or attempted homicide using traditionally available demographic and social variables. If neurocriminology can identify replicable biological risk factors that provide incremental knowledge
Although criminal offending is heterogeneous in nature, a common denominator is that it is immoral. It is conceivable that the neural circuitry underlying moral decision making is impaired in offenders. This moral neural circuit is broadly comprised of the polar and medial prefrontal cortex (PFC), ventral PFC, angular gyrus, posterior cingulate and amygdala. These brain regions have substantial overlap with those regions that are found to be structurally or functionally impaired in offenders\(^1\), (see the figure, which shows a schematic diagram of brain regions that are activated only in moral decision making (green), regions that are impaired only in antisocial groups (red), and regions common to both antisocial behaviour and moral decision making (yellow)). This overlap gives rise to the ‘neuromoral’ hypothesis of antisocial behaviour, which states that some of the brain impairments that are observed in antisocial individuals disrupt moral emotion and/or decision making, thereby predisposing individuals to rule-breaking, antisocial behaviour\(^1\).

This raises an intriguing forensic question. There is little doubt that most violent psychopaths ‘know’ the difference between right and wrong — but do they have the ‘feeling’ of what is right and wrong? Moral decision making is viewed as being influenced by affect\(^1,2\). This ‘moral feeling’, which is centred partly on the amygdala, is argued to be the engine that translates the cognitive recognition that an act is immoral into behavioural inhibition — a mechanism that functions less well in affectively blunted antisocial individuals. Impairments to the emotional component that comprises the feeling of what is moral are viewed as a core feature of psychopaths and are also present in other offenders.

Thus, if a criminal offender has documented disruption to this moral neural circuitry and lacks the feeling for what is right and wrong, are they fully accountable for their immoral behaviour? If this moral circuitry can be better delineated and quantified at the individual level in the future, this affective metric could be entered as a mitigating factor in the punishment phase of a trial, just as low IQ — a cognitive metric — is currently used to establish lack of rational capacity and to excuse the defendant in the guilt phase of a trial.

Intervention and prevention. If neurocriminology could provide even very modest insights into how future offending can be reduced, it would gain considerable traction in the contexts of law and society in general, given that rehabilitation is a consideration in sentencing criminal offenders. Research in this area is currently sparse, but some studies suggest that neurobiological research can inform practice and provide guidelines for future research.

At the psychopharmacological level, it is known from over 45 randomized controlled trials that that a wide range of medications — including atypical antipsychotics, mood stabilizers, stimulants and antidepressants — are effective in reducing aggressive behaviour in children and adolescents\(^3\). Although such effects may in part be due to the treatment of clinical conditions that are co-morbid with aggressive behaviour, such as attention-deficit hyperactivity disorder and depression, pharmacological intervention is also effective in children presenting solely with aggressive symptoms. In adults with impulsive aggression, treatment with selective serotonin reuptake inhibitors has been found to increase glucose metabolism in the orbitofrontal cortex\(^4\), suggesting a potential method for improving functioning in regions that have been identified as deficient in antisocial populations.

Despite these findings, there appear to be few, if any, systematic studies on the long-term efficacy of medications or their application to offender populations. Controversially, anti-androgen medications...
such as medroxyprogesterone or Depo-Provera are thought to reduce recidivism in sex offenders\textsuperscript{107}, but well-controlled randomized controlled trials are lacking. There is agreement that anti-androgens do reduce sexual drive, and in practice at least eight states in the United States have laws on chemical castration. Although some have argued that chemical castration violates the constitutional rights of the offender, others have countered that these medications are effective, that offenders are capable of making an informed decision and that preventing such informed choices that have appropriate safeguards in place is ethically questionable\textsuperscript{108}.

A more socially acceptable avenue of biological intervention may lie in nutritional supplementations such as omega-3 fatty acids. Several studies have documented initial effectiveness in reducing antisocial and aggressive behaviour in child and adult populations\textsuperscript{109–111}, although null findings exist\textsuperscript{112}. The only two randomized controlled trials conducted in prison populations have documented a 34–36% reduction in serious offending in young offenders\textsuperscript{109,110}. Long-chain fatty acids are critical for brain structure and function; they constitute 30% of the cell membrane and are known to enhance neurite outgrowth and prolong cell life\textsuperscript{113}. Given the existence of structural and functional neural correlates of antisocial and violent behaviour and the finding that poor nutrition is an early risk factor for antisocial and aggressive behaviour, omega-3 supplementation may prove to be modestly beneficial for at least some subgroups of offenders.

From a public-health perspective, applications of neurobiological research on violence at the population level relatively early in life may help to prevent adult violence. In one randomized controlled trial, low-income pregnant mothers were provided with prenatal and early postnatal home visitations from nurses who gave advice on reducing smoking and alcohol use and improving nutrition. The study documented a 63% reduction in the number of convictions among the 15-year-old children of these mothers\textsuperscript{114}. One experimental environmental enrichment programme that provided better nutrition, more physical exercise and cognitive stimulation to community children aged 3–5 years documented increased electrocortical arousal and autonomic orienting at the age of 11 years\textsuperscript{115} and a 34.6% reduction in offending rates at the age of 23 years\textsuperscript{116}. In principle, targeted investment of resources to underserved populations at risk of future violence has the potential to enhance neurocognitive functioning and prevent offending, although these initial public health prevention programmes require replication and extension.

Novel, innovative approaches to crime prevention through benign brain manipulation also have the potential to develop from basic neuroscience research. One recent experimental transcranial direct-current stimulation study showed that enhancing neural excitability of the right lateral prefrontal cortex increases compliance to social norms enforced by punishment\textsuperscript{117}. Because crime is a failure to comply with punishment-enforced social norms, and as brain imaging research has documented reduced lateral prefrontal functioning in antisocial groups\textsuperscript{118}, enhancing prefrontal function could, as argued by others, have implications for crime prevention, albeit at a potential cost of reduced compliance to norms that are not sanctioned by punishment\textsuperscript{117}. Mindfulness training has also been experimentally shown to enhance both prefrontal and amygdala functioning\textsuperscript{118,119}, and has been claimed to reduce aggression in offenders\textsuperscript{120,121}. We caution that this potential for crime prevention is extremely preliminary but logically follows from our review of biological risk factors, legal implications and prevention measures. Many would agree that once we can successfully treat offenders, important changes in the law and our social perspective on crime will inevitably ensue.

Conclusions and future directions

Neurocriminological research in particular, and neuroscience in general, are not yet poised to make immediate changes in the prediction, prevention and punishment of criminal offenders. It is also unclear how strong and well replicated scientific findings should be for their proper use in legal cases, although most evidence can be entered as mitigating factors in the penalty phase of a capital punishment case. At the same time, notwithstanding difficulties in determining causality, there is increasing convergence from different disciplinary perspectives that neurobiological influences partly predispose an individual to offending. It is our considered opinion that it would be valuable for researchers and practitioners to focus efforts on: first, the development of innovative and benign biological programmes for crime prevention; second, attempting to enhance the prediction of recidivism, with socially acceptable accuracy, by including neurobiological predictors; third, including emotion alongside cognition in how we legally conceptualize responsibility; fourth, considering the adoption of a dimensional concept of partial responsibility; and fifth, discussing the thorny neuroethical implications of this growing body of neurocriminology research that include the potential for conceptualizing crime as having a physical cause (for example, viewing crime as the result of psychological deficits), stigma and labeling (that is, the potentially harmful effects of identifying individuals based on early biological predispositions\textsuperscript{122}). In conclusion, there is initial proof of concept that neuroscience can become an important future influence in society’s approach to the punishment, prediction and prevention of criminal behaviour.

Andrea L. Glenn is at the Center for the Prevention of Youth Behavior Problems and the Department of Psychology, University of Alabama, Tuscaloosa, Alabama 35487, USA.

Adrian Raine is at the Departments of Criminology, Psychiatry, and Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

Correspondence to A.L.G.
e-mail: Andrea.L.Glenn@ua.edu

doi: 10.1038/nrn6640

Published online 11 December 2015


© 2014 Macmillan Publishers Limited. All rights reserved


Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
Nature Reviews Neuroscience Series on Neuroscience and the law: http://www.nature.com/nrevseries/neurosciellaw/index.html
All links are active in the online PDF

© 2014 Macmillan Publishers Limited. All rights reserved.
Omega-3 supplementation in young offenders: a randomized, stratified, double-blind, placebo-controlled, parallel-group trial

Adrian Raine 1 · Chi-Ching Leung 2 · Melvinder Singh 2 · Jasmin Kaur 2

Published online: 09 January 2020
© Springer Nature B.V. 2020

Abstract

Objectives To examine whether omega-3 supplementation reduces antisocial and aggressive behavior in offenders.

Methods In this randomized, double-blind trial, 145 young offenders were randomized into three groups: omega-3 (N = 48), placebo (N = 46), and treatment-as-usual controls (N = 51). Measures of antisocial, aggressive, and psychopathic behavior were collected at 0 months (baseline), 3 months (end of treatment), 6 months (3 months post-treatment), and 12 months (9 months post-treatment).

Results Omega-3 supplementation resulted in both short-term and long-term declines in self-reported antisocial and aggressive behavior. Findings were stronger for a reactive-impulsive form of aggression than for proactive aggression and psychopathy. Sensitivity analyses documented long-term reductions at 6 and 12 months in the omega-3 group for officer reports.

Conclusions Results suggest that omega-3 supplementation can help reduce antisocial and aggressive behavior over and above regular treatment programs in young offender institutions, particularly for reactive, impulsive aggression.

Keywords Omega-3 · Antisocial · Reactive aggression · Offenders · Psychopathy · Randomized controlled trial

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11292-019-09394-x) contains supplementary material, which is available to authorized users.

✉ Adrian Raine
araine@sas.upenn.edu

1 Departments of Criminology, Psychiatry, and Psychology, University of Pennsylvania, Philadelphia, PA, USA

2 Singapore Prison Service, Singapore, Singapore
Introduction

There is increasing interest in the use of omega-3 as a treatment for aggressive and antisocial behavior. The notion that improved nutrition could reduce antisocial behavior is predicated on risk research documenting that poor nutritional status is a risk factor for externalizing behavior problems and antisocial personality disorder (Liu, Raine, Venables, Dalais, & Mednick, 2004; Neugebauer, Hoek, & Susser, 1999). Omega-3 has been hypothesized as one nutritional component that could explain the link between poor nutrition and antisocial behavior (Raine, Mellingen, Liu, Venables, & Mednick, 2003). Randomized controlled trials (RCTs) have shown some evidence for the efficacy of omega-3 supplementation in reducing antisocial behavior (Raine, Portnoy, Liu, Mahoomed, & Hibbeln, 2015; Raine et al., 2016). Aggressive behavior in particular has been a focus of studies, with a recent meta-analysis of aggression yielding an effect size of $d = 0.24$ (Gajos & Beaver, 2016). As such, there is growing support for a small but significant effect of omega-3 supplementation in reducing antisocial behavior, particularly for aggressive behavior.

From a mechanistic standpoint, there is reason to believe omega-3 could help attenuate neurobiological risk factors for antisocial and aggressive behavior. Animal research has documented that this long-chain fatty acid plays a critical role in brain structure and function, making up approximately 35% of the cell membrane, enhancing neurite outgrowth, and regulating both neurotransmitter functioning and gene expression (McNamara & Carlson, 2006). Structural and functional brain imaging studies on humans have further documented that omega-3 can enhance a variety of brain regions, with no evidence for any detrimental effect (McNamara et al. 2019). Because many of these brain regions overlap with brain areas documented to be impaired in antisocial and violent individuals (Raine 2014), it is reasonable to presume that omega-3 supplementation could help remediate neural risk factors for offending.

Despite progress made in prior studies on omega-3 supplementation for reducing antisocial behavior, two broad questions remain unanswered. The first concerns the form of aggression that omega-3 can attenuate. An important distinction in the literature lies between reactive aggression which is impulsive, emotional, and “hot-blooded,” and proactive aggression which is planned, regulated, and “cold-blooded” (Scarpa, Haden, & Tanaka, 2010). To the extent that omega-3 has been shown to reduce impulsive behavior, omega-3 supplementation may be more efficacious for reactive, impulsive aggression. Only three omega-3 RCT intervention studies have addressed this question. The first from Mauritius documented significant reductions in both forms of aggression (Raine et al., 2015). The second RCT from the USA documented significant reductions in reactive but not proactive aggression (Raine et al., 2016), while the third RCT from Singapore replicated the significant reductions in reactive but not proactive aggression (Raine et al. 2019). Omega-3 supplementation could therefore be expected to be more effective for reactive than proactive aggression, although support to date is limited and mixed.

A second issue concerns the efficacy of omega-3 for offenders. Only two prior prison studies have addressed this issue. One RCT from England provided omega-3 and omega-6 together with multivitamins and minerals for an average of 142 days (range 2 weeks to 9 months). Compared to a placebo group, the experimental group showed a significant 26.3% reduction in prison offenses (Gesch, Hammond, Hampson,
Eves, & Crowder, 2002), although null effects were found for self-report measures (Gesch, 2011). No measures of officer reports were reported. The second RCT conducted in the Netherlands provided omega-3 and multivitamins over an average of 76 days (range 1 to 3 months), and showed a significant reduction in prison incidents, together with a trend for reduced self-report aggression (Zaalberg, Nijman, Bulten, Stroosma, & van der Staak, 2010).

While these two pioneering studies document for the first time some evidence for the efficacy of omega-3 in offenders, they inevitably have some methodological limitations. Gesch et al. (2002) did not include officer-reports, and the dose of DHA and EPA was small (124 mg). In both studies, there was significant variability in treatment duration from prisoner to prisoner, with durations as short as 2 weeks. No treatment-as-usual control groups were included. Because both studies also provided multivitamins, it cannot be concluded that treatment effects were due to omega-3. Importantly, prisoners were not followed up post-treatment to assess whether treatment effects could be sustained over time. Due to the limited sources of antisocial behavior, they were not able to assess for differential effects on different forms of antisocial behavior (e.g., reactive forms of aggression and psychopathic behavior). Drop-out rates in both studies were non-trivial, with rates of 32.2% in Zaalberg et al. (2009) and 25.5% in Gesch et al. (2002). Despite limitations, these two important but initial studies support the hypothesis that nutritional supplementation has efficacy with young offenders.

The present study attempted to address these gaps in the literature and also to assess generalizability of findings from European countries to a south-east Asian sample. The RCT included three groups (treatment-as-usual controls, placebo, and omega-3), with antisocial behavior assessed from self-reports, officer reports, and prison infractions. It was hypothesized that the omega-3 group would show a reduction in antisocial behavior relative to control groups. It was further hypothesized that any treatment effects may be stronger for more reactive, impulsive aggression.

**Methods**

**Trial design**

The design consisted of a randomized, stratified, double-blind, placebo-controlled, parallel-group trial (1:1:1 ratio) of young male offenders. Offenders were randomized into three groups, controls ($n = 51$), placebo ($n = 46$), and omega-3 ($n = 48$), and were assessed at 0 months (start of treatment), 3 months (end of treatment), 6 months (3 months post-treatment), and 12 months (9 months post-treatment). The treatment-as-usual control group was included to assess for any placebo effect as well as the effect of instituting omega-3 supplementation and as such this group was not blinded. Trial design and outcomes remained unchanged throughout the study.

**Participants**

*Eligibility criteria.* Participants had to be aged 16 years or older at the time of study, willing to participate in an RCT, and residing in a reformative training center in
Singapore. Exclusion criteria consisted of (1) allergy to fish or fish products, (2) use of omega-3 supplementation in the past 3 months, (3) intellectual disability, and (4) ineligible to enter reformative training. Written informed consent was obtained from participants and in addition for participants below the age of 18, written parental consent was also obtained, with IRB approval from the University of Pennsylvania.

Study setting, location, and registration. The study took place in a secure institution contained in Changi Prison Complex in Singapore. Participants completed questionnaires in a classroom located within the reformative training center. For officer-rated measures, questionnaires were completed at each officer’s workspace within the institution. The study was registered in ClinicalTrials.gov (NCT03627312) under the title “Omega-3 Supplements to Reduce Antisocial Behavior in Young Offenders” (https://clinicaltrials.gov/ct2/results?cond=&term=SingaporePS&ctrid=&state=&city=&dist).

Omega-3 and standard institutional treatment

Omega-3 supplementation. This consisted of a 200-ml drink (Smartfish Recharge). The base drink in both treatment and placebo conditions consisted of fruit juice from apple, pear, pomegranate, aronia, and passion fruit. It also contained vitamin D (5 μg) and antioxidants (ferric reducing ability of plasma value of 0.71 mmol/100 g). For the treatment condition only, a total of 840 mg of omega-3 (300 mg of DHA, 300 mg of EPA, 180 mg of alpha-linolenic acid, and 60 mg of DPA) was added to the base drink. Placebo drinks were matched exactly with the omega-3 drink in terms of size, appearance, and flavor.

This drink was chosen because (1) it contains an appreciably higher dosage of omega-3 than standard capsules in a relatively small liquid, and (2) a pilot study conducted on the young offenders documented that the fruit-flavored drink was better tolerated and preferred over standard capsules.

Omega-3 treatment duration and administration. Treatment duration was 3 months. The omega-3 drink was administered in the morning before breakfast by the officers on duty.

Omega-3 monitoring. Prison officers monitored the administration of the drinks on a daily basis. They were blind to group assignment and none were part of the research team.

Standard institutional treatment. All three groups received the standard treatment obtained by all young offenders during incarceration. This consisted of academic training, vocational training, religious counseling, a family involvement program, and a community re-integration program. Each inmate was assigned a personal supervisor who provided close guidance and monitoring of the young offender’s welfare and behavior.

Outcome measures

Aggression Questionnaire (AQ) (self-report only). The Aggression Questionnaire (Buss & Warren, 2000) produces five subscale scores for physical aggression, verbal aggression, anger, hostility, and indirect aggression. Extensive support has been documented for convergent, discriminant, and factorial validity, including studies of young offenders (Pechorro, Barroso, Poiares, Oliveira, & Torrealday, 2016).
Youth Psychopathic Traits Inventory (YPI) (self-report only). This instrument (Andershed, Kerr, Stattin, & Levander, 2002) yields three psychopathic personality subscales: Grandiose/Manipulative (interpersonal domain), Callous/Unemotional (affective domain), and Impulsive/Irresponsible (lifestyle/behavioral). Reliability and validity have been documented for this scale in young offenders (Cauffman, Kimonis, Dmitrieva, & Monahan, 2009; Pechorro, da Silva, Rijo, Goncalves, & Andershed, 2017).

Reactive-Proactive Aggression Questionnaire (RPQ) (self-report and officer report). Participants completed this self-report instrument which yields scales of reactive, proactive, and total aggression (Raine et al., 2006). Reliability and validity have been documented (Baker, Raine, Liu, & Jacobson, 2008; Fossati et al., 2009; Cima, Raine, Meesters, & Popma, 2013).

Adult Self-Report (ASR) (self-report only). Participants self-reported on Aggressive Behavior and Rule-Breaking Behavior subscales of the ASR (Achenbach, Rescorla, & Ivanova, 2015). Extensive evidence for reliability, validity, and cross-cultural generalizability has been documented in over 100 societies (Achenbach & Rescorla, 2001).

Conduct and Oppositional Defiant Disorder Scales (CODDS) (self-report and officer report). This 23-item measure is modeled on DSM 5 and assesses the eight DSM oppositional defiant disorder criteria and the 15 conduct disorder criteria (Raine et al., 2016), yielding conduct disorder and oppositional defiant disorder scores. Good construct validity for these scales has been documented (Choy, Raine, Venables, & Farrington, 2017; Raine et al., 2019).

The Antisocial Process Screening Device (APSD) (self-report and officer report). This 20-item scale assesses parent- and child-reported psychopathic traits (Frick, Bodin, & Barry, 2000), yielding three subscales of callous-unemotional, narcissism, and impulsivity. Reliability has been reported as good with the exception of the callous-unemotional scale, with good support for the three-factor model (Salekin, Andershed, & Clark, 2018).

Adult Behavior Checklist (ABCL) (officer report only). Officers rated participants on the Aggressive Behavior and Rule-Breaking Behavior subscales of the ABCL (Achenbach et al., 2015). Extensive evidence for reliability, validity, and cross-cultural generalizability has been documented in over 100 societies (Achenbach & Rescorla, 2001).

Social Dysfunction and Aggression Scale (officer report only). This instrument assesses aggressive behavior as measured by observers (Wistedt et al., 1990). The nine-item Outward Aggression subscale was utilized in this study. Acceptable reliability and construct validity have been documented on institutionalized populations (Wistedt et al., 1990; Grube, 2004).

Institutional infractions. Infractions were measured on a standard list grouped by severity into aggravated (e.g., assault or attack on officer), minor (e.g., quarreling with other prisoners), or warning infractions (similar to minor infractions but where a warning is given).

Dimension reduction. To both provide more robust indices of antisocial behavior and to help reduce type 1 error, all self-report and officer-report scales were separately factor analyzed (see Online supplementary material for full details). This resulted in one overall measure of officer-reported antisocial-aggressive behavior, one overall measure of self-report antisocial-aggressive behavior, and...
three self-report sub-factors (Reactive-Impulsive Aggression, Psychopathy, and Proactive Aggression-Conduct Disorder).

**Blinding success/failure.** Participants in the two drink groups were asked at the end of treatment to guess whether they were in the placebo or omega-3 group.

**Sample size**

Based on prior findings for omega-3 RCTs where the primary outcome was antisocial behavior (Raine, Portnoy, Liu, Mahoomed, & Hibbeln, 2015; Raine et al., 2016), together with results of a meta-analysis of omega-3 RCTs on aggression (Gajos & Beaver, 2016), a small to medium effect size was anticipated. The final total sample size of 145 would have power of 0.80 to detect a small to medium effect size of $f = 0.219$, alpha = 0.05, and critical $F(6,282) = 2.13$.

**Randomization and stratification**

Participants were randomized into treatment, placebo, and control groups by the research coordinator using Urn randomization (Wei, 1978), stratifying on age (four age bands in years—16, 17, 18, 19 and above), ethnicity (Chinese/Malay/Indian/ Others), past violence (yes/no based on current conviction), and violence risk (Low, Low-Moderate, Moderate, Moderate-High, High) based on the Structured Assessment of Violence Risk in Youth (SAVRY—Borum, Bartel, & Forth, 2006).

**Blinding**

All persons involved in data collection and outcome reporting, including participants, research assistants, and prison officers, were blind to the active/placebo omega-3 group allocation (the treatment-as-usual control group was not blinded). Allocation concealment was maintained by having the omega-3 intervention allocation conducted separately by the project coordinator who was kept independent of participants, investigators, and knowledge of which drink codes were omega-3 or placebo. Coding of the drinks was kept only by the first author at an overseas site who had no contact with study participants and was not involved in data collection.

**Statistical methods**

An intention-to-treat (ITT) design using all randomly assigned participants without any exclusion was employed for all data analyses, with data missing due to loss at follow-up handled using mixed effects modeling (Molenberghs & Verbeke, 2005). An ITT approach was adopted because it is considered a gold standard method-of-choice for RCTs, is endorsed by CONSORT, respects initial randomization, and provides unbiased estimates of the effect of treatment assignment on outcome measures (Shrier et al., 2014).

In compliance with CONSORT guidelines (Moher, 2010), the analytic plan focused on documenting group × time interactions. Intervention group, and group × time interaction terms were entered as fixed effects, baseline antisocial score as
a covariate, while outcome measures were modeled using maximum-likelihood estimation with a first-order autoregressive covariance structure and with homogeneous variances to account for the correlation between time points. Counts of institutional infractions which had no missing data and had variance greater than the mean were analyzed using negative binomial regression at each time period. Chi-square analyses were conducted on blinding data (perceived group versus actual group). All tests are two-tailed. All analyses were conducted using SPSS (version 23). False discovery rate control (Benjamini & Hochberg, 1995) was employed to control for type I error on the group comparisons within each time period for each instrument, with uncorrected \( p \) values reported below in the Results section. Figures are based on estimated marginal means. To assess robustness of any significant findings emerging from the ITT design that used mixed-effects models, sensitivity analyses were run using a per protocol design in which only those completing the intervention were analyzed using repeated measures multivariate analysis of variance (RM-MANOVA).

Results

Participant flow, recruitment, and attrition

Full details on participant flow, including enrollment, group allocation, and follow-up, are given in the Online supplementary material. No participant loss was observed on baseline assessment after randomization. Of the 145 participants, 3.4% were lost to follow-up at either 3, 6, or 12 months (one from omega-3, two from placebo, two from controls—see Online supplementary material for detailed reasons for loss). Groups did not significantly differ in this attrition (\( \chi^2 = 0.42, df = 2, p = 0.81 \)).

Demographics and adherence to protocol

Demographics. Demographic data are reported in Table 1. No significant group differences were observed on age, ethnicity, violent/non-violent offending, baseline self-report antisocial behavior, baseline officer-report antisocial behavior, and prison infractions, documenting that randomization and stratification procedures were successful.

Adherence to protocol. Average number of drinks taken per week for placebo and omega-3 groups is provided in Table 1. There was no significant group difference in compliance rates (\( p = 0.72 \)).

Blinding success

Groups did not differ in guessing which of the two drink groups they had been allocated to (\( p = 0.68 \)), with 45.7% of the placebo group believing they were assigned to the omega-3 group, compared to 46.8% in the omega-3 group. This indicates that the blinding was successful.
Adverse events

As with the two prior studies of Gesch et al. (2002) and Zaalberg et al. (2010), no adverse events were reported.

Factor analyses of secondary outcome measures

Self-report and officer-report measures of antisocial behavior were factor analyzed separately to provide overarching indicators and to reduce type I error. Analyses produced single overarching measures of both self-report and officer-report antisocial behavior for each of the four time points. In addition, for self-reports only, three factors consisting of reactive-aggressive, proactive-disruptive, and psychopathy were also extracted. Full details are provided in the Online supplementary material.
Self-Reports

Estimated marginal means from mixed-effects analyses for self-report and officer-report measures and prison infractions at all four time points are detailed in Table 2. Analyses below are based on estimated marginal means.

Self-reports. For the general factor of antisocial behavior, the treatment group × time interaction was significant, $F(6,412) = 2.25, p = 0.038$. Antisocial scores for all three treatment groups are illustrated in Fig. 1a. At 3 months (end of treatment), the omega-3 group were significantly lower than controls ($p = 0.006, d = 0.56$). At 6 months, there was a trend for the omega-3 group to be lower than the placebo group ($p = 0.059, d = 0.39$), which was significant at 12 months ($p = 0.045, d = 0.43$). No other pairwise comparison was significant. The initial group difference at 3 months survived false discovery rate control.

Analyses were conducted on the three factors to assess specificity of findings. For reactive-aggressive scores, the group × time effect was significant, $F(6,408) = 2.18, p = 0.044$ (see Fig. 1b). At 3 months, the omega-3 group was significantly lower than controls ($p = 0.01, d = 0.52$). At 6 months, the omega-3 group was significantly lower than the placebo group ($p = 0.009, d = 0.54$), and they remained significantly lower at 12 months ($p = 0.013, d = 0.51$). No other comparisons were significant. All three comparisons survived false discovery rate control. No group × time interactions were significant for proactive-disruptive ($F = 1.41, p = 0.21$) or for psychopathy factors ($F = 1.29, p = 0.26$).

Prison officer reports. Group effects are illustrated in Fig. 2. The group × time effect was non-significant, $F(6,367) = 0.92, p = 0.48$, although effects were in the same direction as for self-reports. Exploratory post hoc analysis suggested by one reviewer revealed reduced antisocial behavior in the omega-3 group compared to the placebo group at 3 months ($p = 0.019$), 6 months ($p = 0.024$) and also at 12 months ($p = 0.003$).

Prison infractions. Negative binomial regression analyses indicated that group differences were non-significant at all four time points (Wald $\chi^2 = 2.03, df = 2, p > 0.36$).

Sensitivity analyses

Full details of the per protocol robustness analyses are provided in the online supplement.

Self-reports. Significant group × time interactions and the overall pattern of findings were replicated for the general self-report measure of antisocial behavior ($p = 0.045$) and for the reactive-aggressive factor ($p = 0.049$).

Officer reports. A significant group × time interaction was observed for officer reports ($p = 0.046$). Pairwise comparisons indicated that the omega-3 group were lower than the placebo group at 3 months ($p = 0.019, d = 0.52$), 6 months ($p = 0.01, d = 0.56$), and 12 months ($p = 0.001, d = 0.75$), largely paralleling findings from ITT analyses. Comparisons at all time points survived false discovery rate control.

Officer-reported reactive-impulsive aggression. A proxy measure of officer-reported reactive-impulsive aggression produced a significant group × time interaction ($p =$
Table 2. Unadjusted means with 95% confidence intervals (in parentheses) from mixed-effects models on young offenders’ behavioral outcomes in the three intervention groups for the four assessment periods

<table>
<thead>
<tr>
<th></th>
<th>Young offender self-report</th>
<th>Officer report</th>
<th>Prison incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Antisocial</td>
<td>Reactive-aggressive</td>
<td>Proactive-disruptive</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 months</td>
<td>−0.50 (−3.12, 2.12)</td>
<td>−0.16 (−1.61, 1.30)</td>
<td>−0.12 (−1.11, 0.87)</td>
</tr>
<tr>
<td>3 months</td>
<td>2.90 (0.28, 5.52)</td>
<td>1.40 (−0.04, 2.85)</td>
<td>0.89 (−0.10, 1.88)</td>
</tr>
<tr>
<td>6 months</td>
<td>0.41 (−2.23, 3.05)</td>
<td>−0.08 (−1.49, 1.43)</td>
<td>0.18 (−0.82, 1.17)</td>
</tr>
<tr>
<td>12 months</td>
<td>0.08 (−2.57, 2.72)</td>
<td>0.18 (−1.28, 1.64)</td>
<td>−0.28 (−1.28, 0.71)</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 months</td>
<td>0.20 (−2.56, 2.96)</td>
<td>−0.07 (−1.60, 1.45)</td>
<td>0.21 (−0.83, 1.25)</td>
</tr>
<tr>
<td>3 months</td>
<td>−0.66 (−3.44, 2.12)</td>
<td>−0.15 (−1.68, 1.39)</td>
<td>−0.15 (−1.20, 0.90)</td>
</tr>
<tr>
<td>6 months</td>
<td>1.50 (−1.28, 4.28)</td>
<td>1.40 (−0.14, 2.94)</td>
<td>0.33 (−0.72, 1.38)</td>
</tr>
<tr>
<td>12 months</td>
<td>1.93 (−0.88, 4.74)</td>
<td>1.23 (−0.32, 2.79)</td>
<td>0.82 (−0.25, 1.88)</td>
</tr>
<tr>
<td>Omega 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 months</td>
<td>0.34 (−2.36, 3.04)</td>
<td>0.27 (−1.23, 1.76)</td>
<td>−0.09 (−1.11, 0.93)</td>
</tr>
<tr>
<td>3 months</td>
<td>−2.43 (−5.13, 0.27)</td>
<td>−1.35 (−2.84, 0.14)</td>
<td>−0.80 (−1.82, 0.22)</td>
</tr>
<tr>
<td>6 months</td>
<td>−2.23 (−4.94, 0.47)</td>
<td>−1.46 (−2.95, 0.04)</td>
<td>−0.63 (−1.65, 0.39)</td>
</tr>
<tr>
<td>12 months</td>
<td>−2.09 (−4.84, 0.66)</td>
<td>−1.52 (−3.04, −0.00)</td>
<td>−0.51 (−1.55, 0.53)</td>
</tr>
</tbody>
</table>
Main Findings

This study aimed to assess whether omega-3 supplementation could help reduce antisocial and aggressive behavior in young offenders. Partial support was obtained
for supplementation efficacy. Self-reports of overall antisocial behavior were significantly reduced in the omega-3 group both in the short term and the long term. Omega-3 was more effective in reducing reactive, impulsive aggression than proactive aggression or psychopathy both in the short term and the long term. Sensitivity analyses supported the robustness of these self-report findings. While results were non-significant for officer reports based on ITT analyses, post hoc tests nevertheless revealed that the omega-3 groups showed long-term reductions compared to the placebo group at 6 months and 12 months. Sensitivity analyses documented that these significant post hoc officer-report results were buttressed in per-protocol analyses which documented that the omega-3 group evidenced reductions that increased over time in both officer-reported antisocial behavior and also officer-reported reactive-impulsive-aggressive behavior compared to the placebo group, suggesting more robust effects for those completing treatment. No significant reductions in prison infractions were observed. Effect sizes were generally in the low to medium range. Taken together, findings provide partial but not full cross-cultural validation of the two prior prison studies on omega-3 supplementation (Gesch et al., 2001; Zaalberg et al., 2010). It is suggested that further consideration be particularly given to omega-3 supplementation for both offender populations and also non-incarcerated patients presenting with reactive, impulsive forms of aggressive behavior.

**More robust effects for reactive, impulsive forms of aggression**

The most robust effects were found for reactive, impulsive aggression. Importantly, reductions immediately post-treatment at 3 months were sustained at 6 months and 12 months for self-reports at levels which did not diminish over time, lying in contrast to the proactive aggression/disruptive factor which showed no treatment effects. These findings were replicated with officer reports (see Online supplementary material). This finding is consistent with findings from other studies which have reported that reactive
but not proactive aggression is reduced following omega-3 supplementation (Choy & Raine, 2018; Raine et al., 2016, 2019).

What is the mechanism of action whereby omega-3 particularly reduces impulsive, reactive aggression? One possibility is that omega-3 reduced reactive aggression by enhancing the functioning of the prefrontal cortex. Concentrations of DHA vary throughout the brain and are at their highest in the prefrontal cortex (Laye, Nadjar, Joffre, & Bazinet, 2018), an area critical for impulse control and emotion regulation. Higher levels of omega-3 are associated inter alia with increased functional connectivity in the frontal pole and anterior cingulate, areas that subserve executive functions (Talukdar, Zannroziewicz, Zwilling, & Barbey, 2019). Omega-3 supplementation has also been shown to enhance executive functions (McNamara, Asch, Lindquist, & Krikorian, 2018). Given that reactive-impulsive aggression has been associated with reduced glucose metabolism in the prefrontal cortex, as well as poor executive functions (Thomson & Centifanti, 2018) and also reduced connectivity between the prefrontal cortex and amygdala (Romero-Martinez et al., 2019), prefrontal upregulation is a viable explanation for why omega-3 reduces impulsive-aggressive behavior. Future studies could test this hypothesis by evaluating whether omega-3 supplementation enhances prefrontal functioning as assessed by either neurocognitive or brain imaging measures, and by additionally assessing whether such prefrontal upregulation mediates any effect of omega-3 supplementation in reducing antisocial behavior.

**Partial support for omega-3 efficacy**

Findings can only be considered partial support for the notion that omega-3 may have efficacy in reducing general forms of antisocial behavior as no group × time interaction was observed for officer reports using an ITT design incorporating all participants, although post hoc analyses did document significant improvement in the omega-3 group. This caveat must in turn be tempered by three considerations. First, per-protocol analyses of officer reports based on those completing treatment did show a significant group × time interaction, with the omega-3 group compared to the placebo group showing reductions in antisocial behavior post-treatment which increased in size as time progressed. As significant per-protocol effects were also observed for self-reports, important cross-informant corroboration was observed for those completing treatment. Second, officer reports were non-significantly negatively correlated with self-reports ($r = -0.08, p = 0.33$) indicating that they reflect a different form of antisocial behavior. They also failed to yield separable sub-factors that would be expected and which were obtained from self-reports. Third, a prior RCT of omega-3 supplementation on offenders also failed to observe significant effects on all four prison officer measures in the face of trends for self-reported improvement (Zaalberg et al., 2010). These three considerations caution against dismissal of positive findings that emerged and call for future multi-informant research that can resolve these oppositional findings.

A further caveat is the null findings for prison infractions. In contrast, the two prior omega-3 RCTs on offender populations both obtained significant effects for institutional disciplinary records (Gesch et al., 2002; Zaalberg et al., 2010). This discrepancy could be due to the highly skewed distribution of institutional offending in Singapore and relative rarity of serious institutional offending in Singapore compared to Western countries, including the absence of drug and alcohol offenses that have been significant
in other prison studies (Zaalberg et al., 2010). Furthermore, prison infractions were uncorrelated with both officer reports ($r = 0.08$) and self-reports ($r = 0.12$) of antisocial behavior, again indicating the different nature of this prison infraction measure. Recognition of these null findings must be balanced with the significant self-reports and significant officer reports from per-protocol analyses that overall render findings as giving some partial support for the two prior prison studies.

One reason why only partial support was obtained for omega-3 efficacy in reducing antisocial behavior concerns geography. Southeast Asia has the highest omega-3 intake in the world (Micha et al., 2014; Food and Agriculture Organization of the United Nations, 2013), with substantial increases occurring in both seafood and plant omega-3 from 1990 to 2010 (Micha et al. 2014), a time period during which participants in the current study were growing up. Singaporeans have been reported to consume more than five times the seafood intake of those in the United States (United States Department of Agriculture, 2017). Cross-national research has documented a relatively strong negative correlation ($r = -0.68$) between seafood consumption throughout the world and homicide rates, with high seafood consumption associated with low homicide rates (Hibbeln, 2001). Conceivably therefore, there may be a ceiling effect in Singapore of overall higher baseline seafood consumption that limits the impact of omega-3 in enhancing neurobiological functioning and reducing antisocial behavior. This may have detracted from the ability to detect effects with omega-3 supplementation, particularly compared to the two prior studies in England and the Netherlands where omega-3 intake is comparatively lower (Micha et al., 2014). Future research could test whether effect sizes in treatment studies are somewhat larger in countries with less seafood intake than in Singapore.

**Limitations**

Limitations need to be acknowledged. The sample size of 145 is not large, with approximately 48 in each group. While this limits power to detect significant effects, it is twice as large as the median sample size ($N = 22$) calculated from prior omega-3 studies of antisocial behavior (24). Second, while the supplementation period of 3 months is normative for omega-3 studies, longer treatment duration could be more efficacious. Third, while contrasting results as a function of reporter are relatively common in the field (De Los Reyes, Thomas, Goodman, & Kundey, 2013), they underline the need for replication and generalization. Fourth, effect sizes generally were not large, ranging from small to medium, although this is consistent with meta-analytic findings on aggression (Gajos & Beaver, 2016). Fifth, we caution that per-protocol analyses entail missing data that should be treated cautiously compared with ITT analyses which include all participants. These limitations in turn need to be balanced with relative strengths which include the rigorous randomized, stratified, double-blind, placebo-controlled, parallel-group design, the use of multiple report sources that include sub-forms of antisocial behavior, two control groups, low attrition, successful blinding, robustness checks, consistent supplementation for 3 months, and longer-term follow-up 9 months after treatment termination which is rare in this field.
Conclusions

In conclusion, partial support was observed for the efficacy of omega-3 supplementation in reducing antisocial behavior in young offenders, particularly with respect to more reactive, impulsive aggressive behavior. The potential implication of these findings for criminology, psychiatry, and clinical psychology lies in the future potential for a benign neurobiological intervention to reduce antisocial and aggressive behavior in offenders, a treatment approach which is currently not operationalized in prison settings. Given reports of the efficacy of omega-3 supplementation in reducing aggression, anger control, and behavior problems in psychiatric disorders that include borderline personality (Bozzatello, Rocca, & Bellino, 2018), substance abuse (Buydens-Branch & Branchey, 2008), and youth with depression (Young, Arnold, Wolfson, & Fristad, 2017), and given evidence for its efficacy in treating mental disorders (Firth et al., 2019), omega-3 could be particularly effective in reducing antisocial and aggressive behavior in offenders who also present with mental illness.

Acknowledgements This study was made possible with the approval of the Singapore Prison Service. We wish to thank Fann Jiang, Lee Kit Ying, and Doris Chia for their assistance in data collection and data cleaning.

References


**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Behavioral/Cognitive

Stimulation of the Prefrontal Cortex Reduces Intentions to Commit Aggression: A Randomized, Double-Blind, Placebo-Controlled, Stratified, Parallel-Group Trial

Olivia Choy, Adrian Raine, and Roy H. Hamilton

Department of Psychology, Nanyang Technological University, Singapore 637332, Departments of Criminology, Psychiatry, and Psychology, Jerry Lee Center of Criminology, and Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Although prefrontal brain impairments are one of the best-replicated brain imaging findings in relation to aggression, little is known about the causal role of this brain region. This study tests whether stimulating the dorsolateral prefrontal cortex using transcranial direct current stimulation (tDCS) reduces the likelihood of engaging in aggressive acts, and the mechanism underlying this relationship. In a double-blind, stratified, placebo-controlled, parallel-group, randomized trial, 81 human adults (36 males, 45 females) were randomly assigned to an active (N = 39) or placebo (N = 42) condition, and then followed up 1 d after the experiment session. Intentions to commit aggressive acts and behavioral aggression were assessed using hypothetical vignettes and a behavioral task, respectively. The secondary outcome was the perception of the moral wrongfulness of the aggressive acts. Compared with the sham controls, participants who received anodal stimulation reported being less likely to commit physical and sexual assault (p < 0.01). They also judged aggressive acts as more morally wrong (p < 0.05). Perceptions of greater moral wrongfulness regarding the aggressive acts accounted for 31% of the total effect of tDCS on intentions to commit aggression. Results provide experimental evidence that increasing activity in the prefrontal cortex can reduce intentions to commit aggression and enhance perceptions of the moral wrongfulness of the aggressive acts. Findings shed light on the biological underpinnings of aggression and theoretically have the potential to inform future interventions for aggression and violence.

Key words: aggression; moral judgment; prefrontal; transcranial direct current stimulation; violence

Introduction

Prefrontal brain impairment is one of the best-replicated risk factors for aggressive behavior. Evidence from neurological research shows that patients with damage to the frontal cortex exhibit more aggressive behavior (Anderson et al., 1999). In addition to head-injury and lesion studies, the imaging and neuropsychological literature has documented structural and functional prefrontal deficits in antisocial individuals (Brower and Price, 2001; Yang and Raine, 2009). Findings on the role of the frontal cortex in modulating aggression and violence also extend to sexual offending (Chen et al., 2016).

Within the prefrontal cortex, a meta-analysis of 43 imaging studies found that impairments of the dorsolateral prefrontal cortex (DLPFC) are implicated in antisocial behavior, with a stronger effect for the left (d = −0.89) than right (d = −0.56) DLPFC (Yang and Raine, 2009). This may be due to the DLPFC’s broad connection to functions related to aggression, including moral judgment (Mendez, 2009), that can in turn influence the risk of engaging in aggression, a deduction consistent with the neural moral model of antisocial behavior (Raine and Yang,
More recent findings bolster the meta-analytic evidence. The involvement of the DLPFC in aggressive and antisocial behavior has since been documented in other neuroimaging studies (Dalwani et al., 2011; Fairchild et al., 2013; Alegria et al., 2016). Furthermore, while it has been suggested that DLPFC lesions are associated with apathy and diminished motivation (Levy and Dubois, 2006), a meta-analysis of 126 neuropsychological studies measuring executive functions in antisocial populations documented an effect size of $d = 0.44$ for antisocial behavior and $d = 0.41$ for physical aggression, implicating dorsolateral prefrontal dysfunction in aggression (Ogilvie et al., 2011). It is important to recognize, however, that the DLPFC is not the only prefrontal area implicated in antisocial and aggressive behavior. Other sub-regions include the ventromedial prefrontal cortex (Hare et al., 2014) and the anterior cingulate cortex (Kolling et al., 2016), areas which have widespread connections to the DLPFC. Together, studies suggest that there is multimethod evidence indicating the possible implication of the DLPFC on antisocial behavior, among other brain regions. Despite these findings, little is known about the causal role of the prefrontal cortex on aggressive behavior. Conclusions from extant research on the neural foundations of aggression have largely been correlational. Three known studies have tested the effect of prefrontal cortex upregulation on aggression using the Taylor Aggression Paradigm and transcranial direct current stimulation (tDCS), a noninvasive technique that influences neural activity (Brunoni et al., 2012). However, findings have been mixed.

One study documented that upregulating the right DLPFC reduced proactive aggression in males (Dambacher et al., 2015b), while another revealed that increasing left DLPFC activity resulted in more aggressive behavior when participants were angry (Hortenius et al., 2012). In contrast, upregulation of the inferior frontal cortex did not have a significant effect on aggression (Dambacher et al., 2015a). Whether stimulation targeting the DLPFC can reduce intentions to engage in aggressive acts or behavioral aggression using other measures has not been examined and, to our knowledge, no studies have experimentally investigated the intermediary mechanisms linking prefrontal deficits to aggression.

Given the association between prefrontal impairments and aggression, this study tests the hypothesis that upregulating the prefrontal cortex using tDCS will reduce intent to commit an aggressive act. This study additionally extends the limited literature on tDCS and aggression by using a larger sample. As similarities have been found between the neural mechanisms underlying moral cognition in normal individuals and brain mechanisms impaired in antisocial populations (Raine and Yang, 2006), we also assess whether prefrontal upregulation improves judgments of moral wrongfulness, which may in turn partly account for any effect of prefrontal enhancement on reducing intent to commit aggressive acts.

### Materials and Methods

**Trial design.** The study consisted of a double-blind, placebo-controlled, stratified, randomized trial comparing a group that received an anodal tDCS intervention with a sham control group. Baseline assessments and one session of tDCS or sham intervention were conducted during the experimental session, while outcome measures were assessed the following day. Tasks and questionnaires were administered in a fixed order. The study was approved by the Institutional Review Board of the University of Pennsylvania and the trial protocol was registered at ClinicalTrials.gov (NCT02427672).

**Participants.** Eighty-six healthy adults (age, $\geq 18$ years) were recruited in Philadelphia between April 2015 and April 2016. The experiment took place during one visit to the study site. In addition to assessments conducted at baseline, participants were followed up 1 d after the experimental session using a web-based questionnaire. Exclusion criteria included contraindications to brain stimulation, such as metallic implants near the electrode sites; unstable medical conditions; neurological, cardiovascular, or psychiatric illness; participation in another noninvasive brain stimulation study on the same day; history of adverse reactions to tDCS; and lack of e-mail access. Written informed consent was obtained from all participants.

**tDCS intervention.** tDCS was administered by trained study personnel using a battery-driven, constant-current stimulator (TCT Research). Two anodal electrodes were placed over the DLPFC bilaterally (F3 and F4) according to the International 10–20 EEG system. A constant current of 2 mA (1 mA to each DLPFC site) was applied for 2 min through saline-soaked sponge electrodes ($5 \times 5$ cm). A single extracephalic cathodal electrode ($5 \times 7$ cm) was placed at the posterior base of the neck to minimize unintentional effects of inhibitory stimulation on brain activity.

Following standard tDCS protocol, stimulation commenced after a 30 s ramp-up period. The current was ramped down over the last 2 s. The tasks performed during tDCS are understood to influence the behavioral after-effects of stimulation (Gill et al., 2015). Thus, during the stimulation session, all participants performed the Psychology Experiment Building Language (Mueller and Piper, 2014) version of two cognitive tasks known to engage the DLPFC: the Psychomotor Vigilance Task (Dinges and Powell, 1985; Cui et al., 2015), followed by the Iowa Gambling Task (Bechara et al., 1994; Ernst et al., 2002). Although participants in both intervention arms received the same electrode placement and ramp-up/down times, stimulation for the sham control group was discontinued after 30 s. This has proven to be effective for blinding as participants habituate to the sensation of stimulation within seconds of initial initiation (Gandiga et al., 2006).

**Intention to commit aggression.** Behavioral intentions to commit aggressive acts were assessed using two hypothetical vignettes, which have been studied in samples with characteristics similar to ours (Hannon et al., 2000; Mazerolle et al., 2003). Brief scenarios describing two types of aggression, physical assault and sexual assault, were presented to participants, who responded to the anticipated likelihood that they would commit the aggressive act. Responses were measured on a scale ranging from zero (no chance at all) to 10 (100% chance).

**Perceptions of moral wrongfulness.** To assess moral perceptions of the aggressive acts, participants were asked to rate how morally wrong it would be to act as the protagonist in the scenario on a scale from 0 (not at all) to 10 (very). Aggregate measures of aggressive intent and perception of moral wrongfulness were created by combining responses from the physical and sexual assault scenarios (Armstrong and Boutwell, 2012).

**Aggression.** The voodoo doll task is a reliable and validated behavioral analog measure of aggression (Dewall et al., 2013). In this task, participants were shown a computer-based image of a doll that represented a partner or a close friend. They were told that they were given the opportunity to release their negative energy to that individual by inserting as many pins ($0–51$) in the doll as they wished. Instructions did not use the word “voodoo.” Stabbing the doll with more pins indicated higher levels of aggression.

**Randomization and stratification.** At the initial visit, participants were randomized into an active stimulation or sham/placebo condition using a computerized urn randomization procedure (Stout et al., 1994). The stratification factors were age ($18/19/\geq 20$ years), sex (male/female), and ethnicity (Caucasian/non-Caucasian). This stratification was used to balance groups on key demographic variables.

**Blinding.** Participants and experimenters were blind to the tDCS condition assignment. The trial adhered to established procedures to maintain separation between staff that conducted the stimulation and staff that engaged with the participant. In each experimental session, only one experimenter who set up the tDCS procedure had knowledge of the participant’s allocation. To further ensure blinding, all participants were kept blind to the objective of the study and outcome measures were not taken in the presence of research staff as they could lead to biased results.
In the three cases where double blinding was compromised due to the inability of having >1 experimenter at a session, the cases were excluded from analyses. To assess adherence to blinding procedures, James’ (James et al., 1996) and Bang’s (Bang et al., 2004) blinding indices were calculated using the participants’ and blinded experimenters’ guesses about group assignment at the end of the experimental session.

**Statistical analyses.** One-way ANCOVA was used to test group differences in intentions to commit aggression and the behavioral measure of aggression. The following baseline measures were examined as possible covariates: variety of crime throughout the lifetime, aggression, grade point average, trait anxiety, social adversity, psychopathy, the lack of premeditation and sensation-seeking dimensions of impulsivity, and self-control.

In addition to a Self-Report Crime Questionnaire, which asked participants to indicate the number of times they had committed any of 36 criminal and delinquent acts ranging from white-collar and blue-collar offenses (e.g., fraud and shoplifting) to noncriminal, deceptive behaviors (e.g., cheating on an exam), participants’ baseline levels of aggression were assessed using the Reactive–Proactive Aggression Questionnaire (Raine et al., 2006). Trait anxiety was assessed using the 20-item Spielbergser State-Trait Anxiety Inventory (Spielberger, 1983). A social adversity index was obtained based on responses to 14 items obtained from demographic questionnaires. Items indicating adversity included the following: parent unemployment; mother’s low education; father’s low education; parental separation or divorce; placement in a foster home, hospital, or other institution during childhood; having ≥5 siblings; born to a teenage mother; a ratio of people per room (including bedrooms, living room, dining room, and kitchen) of ≥1.0; brought up in public housing; parents’ use of welfare or food stamps from the government; father or mother had been arrested; father or mother has had problems with alcohol or drugs; father or mother has had physical illness, such as heart or lung problems; father or mother has had mental illness, such as alcoholism, major depression, schizophrenia, or anxiety. To assess psychopathic traits, the short form of the Self-Report of Psychopathy-III questionnaire, comprising 29 items, was administered (Paulhus et al., 2009). Additionally, scores were obtained from the lack of premeditation and sensation-seeking subscales of the short-form version of the UPPS-P Impulsivity Scale [D.R. Lynam, “Development of a Short Form of the UPPS-P Impulsive Behavior Scale (2013), unpublished technical report].

Self-control was assessed using the 13-item Brief Self-Control Scale (Tangney et al., 2004).

Following recommendations, stratification variables and baseline measures associated with the outcomes were adjusted for, while variables with baseline imbalances were not [Committee for Proprietary Medicinal Products (CPMP); 2004; Kahan et al., 2014]. Effect sizes were calculated using partial η squared.

To provide information on a mechanism of action accounting for any effect of tDCS on aggressive intent, change in perceptions of moral wrongfulness was examined using ANCOVA. We tested whether the tDCS group perceived aggressive acts as more morally wrong than the sham control group (F(1,70) = 8.40, p < 0.01, η² = 0.11; Fig. 2A). There were no significant interaction effects between treatment group and sex (F(1,70) = 0.57, p > 0.45, η² = 0.01) and between treatment group and ethnicity (F(1,70) = 0.01, p = 0.92, η² < 0.001). Further analyses revealed that intent to commit both physical assault (F(1,70) = 5.61, p = 0.02, η² = 0.07) and sexual assault (F(1,70) = 5.64, p = 0.02, η² = 0.08) were lower in the active tDCS group (Fig. 2A). However, there was no significant group difference in behavioral aggression assessed using the Voodoo doll task (F(1,71) = 1.31, p = 0.26, η² = 0.02; Fig. 2B). Additional sensitivity analysis conducted on log-transformed and square root-transformed data for the aggression measures yielded substantively similar findings (Fig. 2-1, available at https://doi.org/10.1523/JNEUROSCI.3317-17.2018.f2-1).

**Mechanisms accounting for the reduction in intent to commit aggression**

ANCOVA also revealed that compared with controls, the active tDCS group perceived aggressive acts as more morally wrong (F(1,71) = 4.64, p = 0.04, η² = 0.06; Fig. 2C). In particular, the main effect of treatment group was significant for perceptions of moral wrongfulness regarding sexual assault (F(1,71) = 6.81, p = 0.01, η² = 0.09), but not physical assault (F(1,71) = 0.96, p = 0.33, η² = 0.01). Higher ratings of moral wrongfulness partly mediated the reduction in intention to commit aggressive acts (indirect effect: b = −0.51; 95% CI, −1.14 to −0.10; p < 0.05). After controlling for perceptions of moral wrongfulness, treatment group was not a significant predictor of aggressive intent (Fig. 3). Moral perception accounted for 31% of the total effect of treatment group on overall aggressive intent.

Further analysis revealed that moral wrongfulness partly mediated the reduction in likelihood of committing sexual assault (indirect effect: b = −0.34; 95% CI, −1.11 to −0.03; p < 0.05), but not physical assault (indirect effect: b = −0.32; 95% CI, −0.89 to 0.10; p > 0.05). Perceptions of moral wrongfulness accounted for approximately half (PMA = 0.56) of the total effect of treatment group on intent to commit sexual assault. For completeness, sensitivity analyses that included the demographic variables and social adversity as covariates did not substantively change the mediation results (Fig. 3-1, available at https://doi.org/10.1523/JNEUROSCI.3317-17.2018.f3-1).

**Demographics and adherence to protocol**

Baseline distributions of the hypothesized covariates were generally well balanced between the treatment groups. With the exception of social adversity, demographic variables and baseline characteristics did not differ across groups (Table 2). As the James’ blinding indices were >0.5 and Bang’s blinding indices did not approach 1 or −1, participants were considered to have been blinded successfully on average (Table 3; James et al., 1996; Bang et al., 2004).

**Aggression outcomes**

Prognostic covariates were determined based on bivariate associations between the hypothesized covariates and outcome measures (Table 4). A one-way ANCOVA controlling for self-report crime and baseline aggression levels revealed a main effect of treatment group on aggressive intent, with the active tDCS group reporting a significantly lower likelihood of engaging in aggression compared with the sham control group (F(1,70) = 8.40, p < 0.01, η² = 0.11; Fig. 2A). There were no significant interaction effects between treatment group and sex (F(1,70) = 0.57, p > 0.45, η² = 0.01) and between treatment group and ethnicity (F(1,70) = 0.01, p = 0.92, η² < 0.001). Further analyses revealed that intent to commit both physical assault (F(1,70) = 5.61, p = 0.02, η² = 0.07) and sexual assault (F(1,70) = 5.64, p = 0.02, η² = 0.08) were lower in the active tDCS group (Fig. 2A). However, there was no significant group difference in behavioral aggression assessed using the Voodoo doll task (F(1,71) = 1.31, p = 0.26, η² = 0.02; Fig. 2B). Additional sensitivity analysis conducted on log-transformed and square root-transformed data for the aggression measures yielded substantively similar findings (Fig. 2-1, available at https://doi.org/10.1523/JNEUROSCI.3317-17.2018.f2-1).

**Results**

**Participant flow and recruitment**

Data were analyzed on a total sample of 81 (Fig. 1). No participants were lost to follow-up. There was no evidence of selection bias as no significant differences were observed between participants who were included in the analyses and those who were not (p > 0.05; Table 1).

**Adverse events**

tDCS was associated with minimal side effects. No major adverse events were reported over the duration of the study. According to Fertonani et al.’s (2010) scale and consistent with other tDCS...
studies (Brunoni et al., 2012), reported side effects included itchiness (85.2%), lightheadedness (40.7%), pain (46.9%), burning (49.4%), warmth (51.2%), pinching (45.7%), iron taste (7.4%), and fatigue of light-to-moderate intensity (35.0%). No participants withdrew due to these minor events.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart of the screening and enrollment of study participants who were randomly assigned to anodal prefrontal stimulation or a sham control group.

Table 1. Comparison of participants included and excluded in statistical analyses

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Included (n = 81)</th>
<th>Excluded (n = 3)</th>
<th>Statistic</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>45</td>
<td>1</td>
<td>( \chi^2 = 0.58 )</td>
<td>0.45</td>
</tr>
<tr>
<td>Male</td>
<td>36</td>
<td>2</td>
<td>( t = 0.11 )</td>
<td>0.91</td>
</tr>
<tr>
<td>Age</td>
<td>20.21 years (3.31 years)</td>
<td>20.00 years (1.73 years)</td>
<td>( t = 0.11 )</td>
<td>0.91</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>36</td>
<td>1</td>
<td>( \chi^2 = 0.15 )</td>
<td>0.70</td>
</tr>
<tr>
<td>Non-Caucasian</td>
<td>45</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade point average( ^a )</td>
<td>3.59 (0.77)</td>
<td>3.66 (0.29)</td>
<td>( t = -0.17 )</td>
<td>0.87</td>
</tr>
<tr>
<td>Social adversity</td>
<td>1.10 (1.48)</td>
<td>1.00 (1.00)</td>
<td>( t = 0.11 )</td>
<td>0.91</td>
</tr>
<tr>
<td>Variety of offending</td>
<td>16.85 (6.21)</td>
<td>15.00 (5.00)</td>
<td>( t = 0.51 )</td>
<td>0.61</td>
</tr>
<tr>
<td>Baseline aggression</td>
<td>9.37 (4.72)</td>
<td>10.33 (2.08)</td>
<td>( t = -0.35 )</td>
<td>0.73</td>
</tr>
<tr>
<td>Psychopathy</td>
<td>23.17 (12.20)</td>
<td>23.00 (13.75)</td>
<td>( t = 0.02 )</td>
<td>0.98</td>
</tr>
<tr>
<td>Lack of premeditation</td>
<td>1.61 (0.49)</td>
<td>1.92 (0.14)</td>
<td>( t = -1.08 )</td>
<td>0.28</td>
</tr>
<tr>
<td>Sensation-seeking</td>
<td>2.89 (0.64)</td>
<td>2.58 (0.52)</td>
<td>( t = 0.80 )</td>
<td>0.42</td>
</tr>
<tr>
<td>Anxiety</td>
<td>38.60 (8.86)</td>
<td>45.67 (13.05)</td>
<td>( t = -1.34 )</td>
<td>0.19</td>
</tr>
<tr>
<td>Self-control</td>
<td>36.26 (6.95)</td>
<td>37.00 (1.73)</td>
<td>( t = -0.18 )</td>
<td>0.86</td>
</tr>
<tr>
<td>Outcome variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggressive intent</td>
<td>2.26 (3.56)</td>
<td>1.33 (0.58)</td>
<td>( t = 0.45 )</td>
<td>0.66</td>
</tr>
<tr>
<td>Aggression (voodoo doll task)</td>
<td>3.91 (10.29)</td>
<td>3.33 (5.77)</td>
<td>( t = 0.10 )</td>
<td>0.92</td>
</tr>
<tr>
<td>Moral wrongfulness</td>
<td>15.20 (3.48)</td>
<td>16.33 (3.51)</td>
<td>( t = -0.56 )</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table 2. Baseline characteristics by treatment arm

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>tDCS group (n = 39)</th>
<th>Sham group (n = 42)</th>
<th>Statistic( ^a )</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>( \chi^2 = 1.09 )</td>
<td>0.30</td>
</tr>
<tr>
<td>Female</td>
<td>24</td>
<td>21</td>
<td>( t = -0.12 )</td>
<td>0.90</td>
</tr>
<tr>
<td>Male</td>
<td>15</td>
<td>21</td>
<td>( t = -0.12 )</td>
<td>0.90</td>
</tr>
<tr>
<td>Age</td>
<td>20.26 years (4.13 years)</td>
<td>20.17 years (2.36 years)</td>
<td>( t = -0.12 )</td>
<td>0.90</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>( \chi^2 = 0.02 )</td>
<td>0.88</td>
</tr>
<tr>
<td>Caucasian</td>
<td>17</td>
<td>19</td>
<td>( t = -1.18 )</td>
<td>0.24</td>
</tr>
<tr>
<td>Non-Caucasian</td>
<td>22</td>
<td>23</td>
<td>( t = -1.18 )</td>
<td>0.24</td>
</tr>
<tr>
<td>Grade point average( ^c )</td>
<td>3.55 (0.27)</td>
<td>3.47 (0.33)</td>
<td>( t = 1.15 )</td>
<td>0.22</td>
</tr>
<tr>
<td>Social adversity</td>
<td>0.72 (1.15)</td>
<td>1.45 (1.67)</td>
<td>( t = 2.32 )</td>
<td>0.02</td>
</tr>
<tr>
<td>Variety of offending</td>
<td>17.36 (6.25)</td>
<td>16.38 (6.22)</td>
<td>( t = -0.71 )</td>
<td>0.48</td>
</tr>
<tr>
<td>Baseline aggression</td>
<td>9.92 (4.97)</td>
<td>8.86 (4.48)</td>
<td>( t = -1.02 )</td>
<td>0.31</td>
</tr>
<tr>
<td>Psychopathy</td>
<td>23.33 (11.85)</td>
<td>23.02 (12.65)</td>
<td>( t = -0.11 )</td>
<td>0.91</td>
</tr>
<tr>
<td>Lack of premeditation</td>
<td>1.59 (0.49)</td>
<td>1.63 (0.49)</td>
<td>( t = 0.38 )</td>
<td>0.71</td>
</tr>
<tr>
<td>Sensation-seeking</td>
<td>2.89 (0.65)</td>
<td>2.88 (0.65)</td>
<td>( t = 0.07 )</td>
<td>0.94</td>
</tr>
<tr>
<td>Anxiety</td>
<td>38.79 (8.53)</td>
<td>38.43 (9.25)</td>
<td>( t = 0.19 )</td>
<td>0.85</td>
</tr>
<tr>
<td>Self-control</td>
<td>37.05 (6.69)</td>
<td>35.52 (7.18)</td>
<td>( t = 0.99 )</td>
<td>0.33</td>
</tr>
</tbody>
</table>

\( ^a \)Data for continuous variables are presented as mean (SD).
\( ^b \)Differences in baseline scores were compared using two-tailed independent t test and \( \chi^2 \) tests.
\( ^c \)For eight individuals whose grade point averages were missing, mean imputation was conducted. Missing values were replaced with the mean of the observed data as suggested in Kahan et al. (2014).
Table 3. Participant and experimenter conjectures about group assignment and blinding indices

<table>
<thead>
<tr>
<th>Intervention</th>
<th>tDCS</th>
<th>Sham</th>
<th>Do not know</th>
<th>Total</th>
<th>James’ blinding index</th>
<th>Bang’s blinding index</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>tDCS</td>
<td>26 (32.1)</td>
<td>3 (3.7)</td>
<td>10 (12.3)</td>
<td>39 (48.1)</td>
<td>0.59</td>
<td>0.42, 0.76</td>
<td></td>
</tr>
<tr>
<td>Sham</td>
<td>18 (22.2)</td>
<td>9 (11.1)</td>
<td>15 (18.5)</td>
<td>42 (51.9)</td>
<td>-0.21</td>
<td>-0.41, -0.02</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44 (54.3)</td>
<td>12 (14.8)</td>
<td>25 (30.9)</td>
<td>81 (100)</td>
<td>0.57</td>
<td>0.49, 0.65</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimenter’s guess, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tDCS</td>
</tr>
<tr>
<td>Sham</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

*Due to missing data, three cases were omitted from calculations of the blinding indices.

Table 4. Relationships between outcome variables (aggressive intent, moral wrongfulness, behavioral aggression) and baseline characteristics of the sample, assessed using t tests for dichotomous demographic variables (upper section) and Pearson correlations for continuous baseline variables (lower section)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Aggressive intent</th>
<th>Moral wrongfulness</th>
<th>Behavioral aggression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>−2.10*</td>
<td>4.21***</td>
<td>0.11</td>
</tr>
<tr>
<td>Race</td>
<td>−0.08</td>
<td>0.06</td>
<td>−0.52</td>
</tr>
<tr>
<td>Age</td>
<td>−0.07</td>
<td>−0.01</td>
<td>−0.02</td>
</tr>
<tr>
<td>Grade point average</td>
<td>0.13</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td>Social adversity</td>
<td>−0.08</td>
<td>0.09</td>
<td>−0.05</td>
</tr>
<tr>
<td>Variety of offending</td>
<td>0.36***</td>
<td>−0.21</td>
<td>0.001</td>
</tr>
<tr>
<td>Aggression</td>
<td>0.42***</td>
<td>−0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Psychopathy</td>
<td>0.17</td>
<td>−0.30**</td>
<td>0.20</td>
</tr>
<tr>
<td>Lack of preméditation</td>
<td>−0.07</td>
<td>0.11</td>
<td>0.28</td>
</tr>
<tr>
<td>Sensation-seeking</td>
<td>0.17</td>
<td>−0.06</td>
<td>0.19</td>
</tr>
<tr>
<td>Anxiety</td>
<td>−0.02</td>
<td>−0.07</td>
<td>0.22</td>
</tr>
<tr>
<td>Self-control</td>
<td>0.01</td>
<td>−0.07</td>
<td>0.22</td>
</tr>
</tbody>
</table>

*p < 0.05; **p < 0.01; ***p < 0.001.

Discussion

This study tested a new approach to reducing aggressive and violent behavior. Individuals who underwent bilateral anodal stimulation of the DLPFC using tDCS reported a lower likelihood of committing an aggressive physical and sexual assault 1 d after stimulation compared with a sham control group. The treatment–aggressive intent relationship was partly accounted for by enhanced perception that the aggressive acts were more morally wrong, resulting from prefrontal upregulation. Findings help to strengthen conclusions from neurological, neuroimaging, and neuropsychological research (Damasio et al., 1994; Damasio, 2000; Yang and Raine, 2009; Liljegren et al., 2015; Rogers and De Brito, 2016) by documenting experimentally the role of the prefrontal cortex on the likelihood of engaging in aggression and the perception of such acts as morally wrong.

Beyond examining experimentally the role of the prefrontal cortex on a behavioral symptom, the finding that moral judgment partly mediates the effect of tDCS on the likelihood of sexual assault contributes to our mechanistic understanding of the etiology of sexual violence. It also provides partial support for the neuro–moral theory of violent behavior, which postulates that violence is due in part to impairments in brain regions subserving moral cognition and emotion (Raine and Yang, 2006). The null mediation effect observed for physical assault suggests that moral judgment plays a greater role on intentions to commit sexual assault, which is consistent with empirical evidence that sexual offenses, such as rape, are rated as more morally wrong than physical violence (Akman et al., 1968; Hsu, 1973). This indicates that moral judgment is likely only one of several processes underlying the prefrontal–aggression relationship.

The difference in our results for behavioral intent and the behavioral measure of aggression warrant attention. Although participants in the tDCS group exhibited significantly lower levels of aggressive intent after the experimental session, they exhibited a nonsignificant increase (d = 0.26) in behavioral aggression. These null findings converge with the mixed findings on tDCS and behavioral aggression in the literature to date (Hortensius et al., 2012). Furthermore, a recent case study of two female patients receiving anodal tDCS over the left DLPFC and a cathode over the right DLPFC reported anger attacks after stimulation, although notably, in contrast to the present study, these subjects were diagnosed with major depressive disorder (Hung and Huang, 2017).

Given empirical evidence that changes in intentions precede behavioral change (Webb and Sheeran, 2006), our results indicating lower intent to engage in aggressive acts following anodal prefrontal stimulation suggest that tDCS may be an initial step toward the reduction of aggression. This implication must, however, be tempered with the mixed findings in the extant literature. While the treatment and control groups did not differ on the behavioral measure of aggression, this finding is consistent with the concept that a single session of tDCS may have a limited effect on behavioral change. The longer-lasting therapeutic effects of tDCS are suggested to be associated with repeated, rather than single, sessions of stimulation (Nitsche et al., 2008). Therefore, beyond intent to engage in aggression, future studies need to evaluate whether behavioral changes may be observed with more stimulation sessions.

Several caveats are in order. First, the trial findings are limited to an ostensibly healthy population. As the first study to test the effect of prefrontal cortical upregulation on aggressive intentions, the generalizability of the findings to other samples remains to be seen. A second limitation is that moral judgment and aggressive intent were measured concurrently. Thus, we were unable to confirm the temporal order of the mediator and outcome variable. However, empirical evidence that moral judgments shape behavior (Reynolds and Ceramic, 2007) supports the notion that the mediation model presented reflects the expected temporal effects. Third, this study measured aggressive inclinations 1 d after the intervention. Further research is needed to determine whether tDCS can produce longer-term reductions in aggressive intent, as well as any reduction in aggressive behavior. Fourth, we were not able in our design to include stimulation of a “control” brain region to help document specificity of findings to the DLPFC. Although it has been documented that the right DLPFC is involved in moral judgment (Tassy et al., 2012), this study did not consider any laterality effects. Fifth, although the findings demonstrate that anodal tDCS resulting in a current flow through the
Figure 2.  A–C, Group means for (A) aggressive intent, (B) behavioral aggression, and (C) perceptions of moral wrongfulness at follow-up. *p < .05, **p < .01. Extended data are presented in Figure 2-1, available at https://doi.org/10.1523/JNEUROSCI.3317-17.2018.f2-1.
DLFPC influences intentions to commit aggression, they do not negate the involvement of other prefrontal areas, such as the ventromedial and anterior prefrontal cortex, or of nonprefrontal areas, including the temporoparietal cortex. Future studies using complementary noninvasive neurostimulation approaches, such as transcranial magnetic stimulation and high-definition tDCS, may elucidate the anatomical specificity of this effect and the complexity of the functional neuroanatomy of violent behavior.

There has been increasing discussion of biological interventions on antisocial and aggressive behavior in both children and adults (Gesch et al., 2002; Raine et al., 2015; Hübner and White, 2016). Our initial findings, which are limited to intentions to commit aggression and moral judgment, require extensive replication. Nevertheless, among other etiological mechanisms, the role of biological factors on the development of antisocial behavior, including aggression, has been increasingly acknowledged (Raine, 2002; Glenn and Raine, 2014; Latvala et al., 2015). It has been suggested that treatment programs will be improved by considering biological mechanisms that potentially regulate aggression (Beauchaine et al., 2008). Thus, it can be argued that further investigation of basic science trials on tDCS may potentially offer a promising new biological approach for reducing aggression, which is a major public health problem and a feature of a variety of mental disorders, including antisocial personality disorder, intermittent explosive disorder, conduct disorder, and borderline personality disorder (American Psychiatric Association, 2013).

Conclusion

Understanding the etiology of aggression and the development of new interventions are paramount to a public health approach to violence reduction (Butchart et al., 2004; Slatkin, 2017). This first known application of prefrontal tDCS to study intentions to commit aggression takes a modest step toward advancing knowledge about the neural mechanisms that regulate aggression. Findings provide experimental evidence for the role of the prefrontal cortex on both physical and sexual assault, and suggest how the brain may, in theory, be amenable to change using a noninvasive tool with transient and relatively minor adverse effects (Poreisz et al., 2007; Fertonani et al., 2015). Nevertheless, a stronger evidence base that includes more consistent findings, documentation of long-term beneficial effects, and a comprehensive effort to rule out potentially aversive side effects is required before this technique can be considered in practice to reduce the perpetration of aggressive acts.

References


Chen CY, Raine A, Chou KH, Chen YJ, Hung D, Lin CP (2016) Abnormal white matter integrity in rapists as indicated by diffusion tensor imaging. BMC Neurosci 17:45. CrossRef Medline


Dewall CN, Finkel EJ, Lambert NM, Sloter EB, Boddenhausen GV, Pond RS Jr,


Hare TA, Hakimi S, Rangel A (2014) Activity in dlPFC and its effective connectivity to vmPFC are associated with temporal discounting. Front Neurosci 8:50. CrossRef Medline


Sulkin G (2017) Reducing violence as the next great public health achievement. Nat Hum Behav 1:0025. CrossRef


Using Taint Analysis and Reinforcement Learning (TARL) to Repair Autonomous Robot Software

D. Lyons and S. Zahra
Robotics & Computer Vision Laboratory
Dept. of Computer & Information Science
Fordham University
New York 10458 USA

Abstract—It is important to be able to establish formal performance bounds for autonomous systems. However, formal verification techniques require a model of the environment in which the system operates; a challenge for autonomous systems, especially those expected to operate over longer timescales. This paper describes work in progress to automate the monitor and repair of ROS-based autonomous robot software written for an a-priori partially known and possibly incorrect environment model. A taint analysis method is used to automatically extract the data-flow sequence from input topic to publish topic, and instrument that code. A unique reinforcement learning approximation of MDP utility is calculated, an empirical and non-invasive characterization of the inherent objectives of the software designers. By comparing off-line (a-priori) utility with on-line (deployed system) utility, we show, using a small but real ROS example, that it’s possible to monitor a performance criterion and relate violations of the criterion to parts of the software. The software is then patched using automated software repair techniques and evaluated against the original off-line utility.

Keywords—self-healing, autonomous system, performance guarantee, automated repair, ROS, reinforcement learning

I. INTRODUCTION

As robot systems are increasingly deployed in domestic, industrial and military applications, it is increasingly important to develop a method to establish bounds for how systems will perform under realistic operating conditions. While there is an established engineering tradition of using mathematics to model physical structures and processes sufficiently to guarantee that a construction will fulfill its specification, there are challenges applying this to autonomous systems [1]. Not many software developers have the skill set necessary to formally establish program correctness according to a specification. But even if they did, formal verification of software [2] [3] and autonomous systems [4], as well as ‘correct by construction’ formal synthesis methods [5] all suffer from the problem that they require a designer-specified model of the environment. And that model can be incomplete because of designer oversight or oversimplification of a complex and dynamic environment. For autonomous systems that are expected to operate over longer timescales, there is also the issue of encountering significant drift over time between the model and the environment encountered by the agent.

Inspired by progress in empirical automated software repair [6] we propose an approach to the automatic analysis and repair of autonomous robot system software to handle environment model related performance failures. Leveraging static analysis techniques [7] used in compilers and security analysis, the approach does not require software engineers to do more than they are doing now in terms of coding. It aims to accomplish for autonomous systems, what approaches like SapFix [8], Clearview [9] and others [6] have achieved in automated software patching.

The remainder of the paper is laid out as follows: Section II introduces the basic ideas in our approach. Section III uses a short example to illustrate our non-invasive approach to analysis and repair. Section IV summarizes the direction of the work and the issues of scaling to future plans.

II. TARL APPROACH

The first assumption of our proposed approach is that while most autonomous robot software is not developed using formal verification techniques, nonetheless, software engineering experience is such that the systems produced by experienced software developers are a solid first approximation to a system with formal performance guarantees. We will begin our approach to formal performance guarantees by leveraging the skills of software developers. This has the advantages of not asking the software development community to adopt new formalisms, and widening the scope of our method, at least in principal, to include any existing (ROS) robot software.

Our second assumption is that experienced software professionals are capable of writing a performance monitor that will signal when their software is working as they intended it to work (i.e., to their apprehension of the specification). The field of runtime verification addresses this capability formally [10], but our need is for far less; an informal monitor would suffice.

It is extremely common for the developers of autonomous robot software to test their algorithms using a simulation tool [11]. While it is true that simulation tools do not replace physical testing, especially for uncertain environments, they nonetheless have become very powerful and comprehensive – driven often by the needs of realistic gaming applications. It is very likely still true that “Simulation is doomed to succeed” [12] in the sense that simulations only model what the designer expects to
 happen (or v.v.). However, software designers’ use of simulation opens for us an approach to non-invasive and automatic formal characterization of working software, for at least the conditions that the simulation implements. And we will show that ultimately this will not be a restriction.

A. Formal Model

An autonomous robot system operating in a partially known environment will be modeled as an MDP $M(S, A, P, R)$ where the state $S$ is a function of sensor signals (we will augment this presently), the actions $A$ are the actuator commands, the transition function $P$ is the stochastic environment, and reward $R$ is the feedback on whether the system is satisfying its performance specification. A policy $\pi$ for $M$ selects which action to carry out based on the state. We will consider the human-written system software to implicitly embody the policy for $M$, and the human-written performance monitor to implicitly embody the reward $R$.

The utility $U(s) s \in S$ for an MDP is a measure of the long-term desirability of each $s \in S$ and can be written via a Bellman eq:

$$U(s) = R(s) + \gamma \sum_{s'} P(s \mid s, \pi(s)) U(s')$$

for discount $\gamma$ and deterministic policy $\pi$. We propose that utility presents a non-invasive and automatic way to characterize the performance of autonomous system software, to indicate where the software may be lacking when the performance monitor fails, and to automatically repair the software. Reinforcement learning [13] gives us a way to calculate utility, but only if the software can be automatically instrumented for the calculation.

B. ROS

The ROS robot operating system is extremely widely used and has greatly improved the portability of robot code since its release in 2007. A ROS system is composed of computing nodes that communicate over named channels called topics. Sensor driver software publishes incoming measurements on sensor topics. Other computing nodes can register callback functions that will be triggered whenever data arrives on a topic. Actuator drivers, that control the mechanism, register callbacks that are triggered whenever data is written to actuator topics. We will restrict our attention to ROS software both because of its popularity and its convenient topic/node structure. A ROS computation graph shows computation nodes connected via the topics they communicate on. This makes it very clear where data flows are coming in and going out.

C. Taint Analysis and Reinforcement Learning (TARL)

The performance monitor $R$, written by the human software designers, tests all the input topics necessary to determine if the performance guarantee holds. It is necessary to find an automatic way to identify the connections, line by line through computing node source code, between input topics and actuator topic publishing commands. While some of this data flow is evident in the ROS computation graph, most of it is hidden within the source code. We will automatically find these data flows and instrument them so that the utility can be estimated in terms of sensor state and statement number. We will then show how this information can be used to detect when software is no longer performing correctly due to unanticipated aspects of the environment, and to potentially repair it.

Taint analysis [7] is a static analysis technique employed in compiler optimization and security analysis. The analysis starts with a program and storage location to investigate for a potential attack, the sink. Any user input location is a taint source. The program is investigated to see what additional storage locations become tainted in storing any computations that involve a tainted location and whether this leads to the sink. The sequence of statements between source and sink is the taint list. We use PyT [14] (modified for ROS) to find taint lists in ROS/Python code and instrument them for utility learning.

```
20 def callback(data): # collect position sensor
21    pos = data.pose.pose.position # from topic
22
23 def travel(goal,vout): # move to goal
24    vel = 5 * delta
25    err = abs(delta)
26    if __name__=='__main__':
27    try:
28       travel(goal,vout)
29       travel(G2,vpub)
30       while True: # go back and forth
31       travel(G1,vpub)
32       travel(G2,vpub)
33       except rospy.ROSInterruptException:
34
Taint List:

(‘if__name__==‘__main__’’, 32)
(‘data=Odometry’, 35)
(‘pos=data.pose.pose.position’,22)
(‘while True’, 37)
(‘delta=goal- pos’,27)
(‘vel=5*delta’,29)
(‘vout.publish(vel)’,30)
```

Figure 1: ROS/Python program implementing a recurrent robot task (top); output of ROS-modified PyT taint analysis (bottom).
flow value. This last unique addition to the SARSA rule makes utility sensitive not just to control-flow, but also to how the information is being manipulated at each line.

III. EXAMPLE

Figure 1(top) shows python code for a ROS node that shuttles a robot between two goal locations (e.g. delivery robot). The performance monitor (not shown) is that the robot reaches each goal within a time and precision bound. The taint list (Fig. 1(bottom)) tags the data flow in the program and is instrumented for RL estimation of utility as described. For convenience, the data below shows just the G1 to G2 transit.

A. Offline Utility

Let us consider that the robot has an odometry sensor and velocity output, and also a binary terrain type sensor. This latter is unused in Fig.1 as the designers consider the terrain won’t affect the performance. Fig. 2 shows the utility calculated by the SARSA algorithm converged for the performance specification (reward) for the ‘off-line’ case, i.e., before the software is deployed. The red points are the terrain sensor true case, and blue are the false case. Line number are the left axis and odometry the right. This graph is an empirical characterization of the software designers’ intent, subject to perhaps faulty information about the environment.

![Figure 2: Utility table run under expected environment conditions](image)

B. On-line Utility

Now let us consider what happens when our autonomous robot system software is deployed, and it is found that the environment does not behave as expected. It is found that muddy terrain conditions are sometimes experienced in transit, which slows down the progress of the robot towards its G2 goal. Using the same RL framework, utility from real life missions is collected; some missions experience the terrain issues and (sometimes) fail and some experience the same environment the designers expected. In this paper, to simplify the time aspects of testing, we don’t run a physical robot mission. We run another version of the off-line simulation in which the environment model is changed.

Fig. 3 shows the utility collected during missions which sometimes failed because of unexpected and unmodeled terrain conditions. The horizontal axis is odometry and the colors are as before. Looking at the difference of off-line and on-line utilities we can isolate a region of maximum utility difference, indicated in Fig 3 by the dashed rectangle. In that region we calculate the average utility curves for the taint list on-line and off-line cases and compare them

\[
q_{\pi}(n) = \sum_{m, p \in \text{diff}} q_{\pi}(m, p, n)
\]

Where \( n \) is line number, \( m \) is terrain sensor, \( p \) is odometry and \( q_{\pi} \) is the utility.

![Figure 3: Utility table under unexpected environment conditions](image)

Figure 4(a) shows this comparison, which evidences a peak at the 2\textsuperscript{nd} last line (vel=5*delta). As it shows the maximum change in utility, we propose this identifies it as the most likely culprit for the failure and, hence, the candidate for repair analysis.

![Figure 4: Comparison of off-line (blue) and on-line (yellow) utility slices \( q_{\pi} \) with difference (red) (a) original, (b) lines swapped.](image)

It is reasonable to ask if the peak simply precedes the line closest to the reward (the publish). To show that our addition of flow information to the SARSA rule addresses this, the same test was done but with Fig 1 lines 29 and 30 swapped. The result is shown in Fig. 4(b) which now has a peak at the 3\textsuperscript{rd} last line, correctly reflecting the swap. The addition of the reward-scaled flow value to the SARSA update equation is what makes utility sensitive to the values being calculated.

C. Automated Repair

Automated program repair [6] is an area of increasing interest to automate and speed up the patching of commercial software bugs, especially security related bugs and for self-healing autonomous devices. The process is typically divided into two phases: determining where the bug might be located in the software and proposing/testing fixes to the bug. Our approach here will address the single line culprit model we developed in the prior section. However, there is nothing that would limit this being applied to a multiple line culprit model. We will also continue to use RL as a solution method – this time to search
over the space of program modifications. Determining how to edit the code will be divided into two steps:

1. Mutation of the affected code to generate potential software patches
2. A modified $\varepsilon$-Greedy policy and SARSA TD-RL algorithm to determine which mutation produces optimal performance

We use the Python AST library to determine any constants in the identified line and generate a set of mutated values for these constants [8]. Every copy of the mutated instruction is protected (i.e., within an IF statement) by the condition used to make the value table slices, Eq. (2). Rather than testing one mutation at a time, all mutations are tested by selecting among them with an $\varepsilon$-Greedy policy and monitoring average total reward ($ATR$) compared to that of the initial, off-line performance

$$ATR = \frac{\sum_t R_t}{E}$$

where $E$ is the number of episodes and $\sum_t R_t$ the total reward up to time $t$. Fig. 5 shows a graph of $ATR$ for the off-line case (blue), the on-line case that encounters problems (yellow), the $\varepsilon$-Greedy search of mutations for repair (green), and the final selected repair (red). The $ATR$ performance of the $\varepsilon$-Greedy search, while it is better than the on-line performance, is not as good as the original off-line performance. The reason is that the $\varepsilon$-Greedy modification conducts significant search exploration of the mutated code, picking up a lot of performance hits. The curve is slowly rising and looks as if it will approach the offline value eventually. However, once we estimate the optimal mutation, we can simply re-estimate the utility using only that optimal mutation. That is the function shown in red on the graph, re-establishing the performance guarantee.

![Figure 5: Code mutation and evaluation for the running example](image)

IV. CONCLUSIONS

While it is becoming evident that its necessary to establish performance guarantees for autonomous systems, this is very challenging in practice. Software designers rarely have the mathematical skills to use significant formal verification. And even if they did, uncertainty about the environment encountered in execution (their specification for verification) would weaken this approach in any case. We are inspired by the extremely practical work done in automated software repair [8] to propose an automated analysis and repair approach for existing software for autonomous robot systems.

This paper describes work in progress, and the results presented are from a thorough but small feasibility study. Our next steps are to apply this to a large quadrotor mission in ROS-M. Issues of scale need to be addressed in a mission with a multipart performance guarantee resulting in multiple taint lists of varying lengths being identified. By calculating one utility per taint list and using the full multi-part performance monitor as reward for every list, we propose to avoid the potential exponential increase in utility size ($2^{\text{num}_\text{crit}}$) while allowing for performance related cross-effects between taint flows. The example here uses a tabular utility. We expect however to need topic-specific, parametric forms in future work.

V. REFERENCES

Consensus statement on abusive head trauma in infants and young children

Arabinda Kumar Choudhary 1 · Sabah Servaes 2 · Thomas L. Slovis 3 · Vincent J. Palusci 4 · Gary L. Hedlund 5 · Sandeep K. Narang 6 · Joëlle Anne Moreno 7 · Mark S. Dias 8 · Cindy W. Christian 9 · Marvin D. Nelson Jr 10 · V. Michelle Silvera 11 · Susan Palasis 12 · Maria Raissaki 13 · Andrea Rossi 14 · Amaka C. Offiah 15

Received: 16 November 2017 / Revised: 22 March 2018 / Accepted: 25 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract
Abusive head trauma (AHT) is the leading cause of fatal head injuries in children younger than 2 years. A multidisciplinary team bases this diagnosis on history, physical examination, imaging and laboratory findings. Because the etiology of the injury is multifactorial (shaking, shaking and impact, impact, etc.) the current best and inclusive term is AHT. There is no controversy concerning the medical validity of the existence of AHT, with multiple components including subdural hematoma, intracranial and spinal changes, complex retinal hemorrhages, and rib and other fractures that are inconsistent with the provided mechanism of trauma. The workup must exclude medical diseases that can mimic AHT. However, the courtroom has become a forum for speculative theories that cannot be reconciled with generally accepted medical literature. There is no reliable medical evidence that the following processes are causative in the constellation of injuries of AHT: cerebral sinovenous thrombosis, hypoxic–

Thomas L. Slovis passed away before publication of this work was completed.

Arabinda Kumar Choudhary
ac0026@nemours.org

1 Department of Radiology, Nemours AI duPont Hospital for Children, 1600 Rockland Road, Wilmington, DE 19803, USA
2 Department of Radiology, The Children’s Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, USA
3 Department of Radiology, Children’s Hospital of Michigan, Wayne State University, Detroit, MI, USA
4 New York University School of Medicine, New York, NY, USA
5 Department of Medical Imaging, Primary Children’s Hospital, Intermountain Healthcare, Department of Radiology, University of Utah School of Medicine, Salt Lake City, UT, USA
6 Division of Child Abuse Pediatrics, Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, IL, USA
7 Florida International University College of Law, Miami, FL, USA
8 Departments of Neurosurgery and Pediatrics, Penn State Health Children’s Hospital, Hershey, PA, USA
9 Department of Pediatrics, Child Abuse and Neglect Prevention, The Children’s Hospital of Philadelphia, The Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA
10 Department of Radiology, Children’s Hospital of Los Angeles, Los Angeles, CA, USA
11 Department of Radiology, Boston Children’s Hospital, Boston, MA, USA
12 Pediatric Neuroradiology, Children’s Healthcare of Atlanta, Scottish Rite Campus, Department of Radiology, Emory University School of Medicine, Atlanta, GA, USA
13 Department of Radiology, University Hospital of Heraklion, University of Crete, Heraklion, Crete, Greece
14 Neuroradiology Unit, Istituto Giannina Gaslini, Genoa, Italy
15 Paediatric Musculoskeletal Imaging, Academic Unit of Child Health, Sheffield Children’s NHS Foundation Trust, Western Bank, University of Sheffield, Sheffield, UK

Published online: 23 May 2018
ischemic injury, lumbar puncture or dysphagic choking/vomiting. There is no substantiation, at a time remote from birth, that an asymptomatic birth-related subdural hemorrhage can result in rebleeding and sudden collapse. Further, a diagnosis of AHT is a medical conclusion, not a legal determination of the intent of the perpetrator or a diagnosis of murder. We hope that this consensus document reduces confusion by recommending to judges and jurors the tools necessary to distinguish genuine evidence-based opinions of the relevant medical community from legal arguments or etiological speculations that are unwarranted by the clinical findings, medical evidence and evidence-based literature.

**Keywords**  Abusive head trauma - Child abuse - Children - Computed tomography - Consensus statement - Infants - Magnetic resonance imaging - Mimics - Unsubstantiated theories

**Executive summary**

This consensus statement, supported by the Society for Pediatric Radiology (SPR), European Society of Paediatric Radiology (ESPR), American Society of Pediatric Neuroradiology (ASPRN), American Academy of Pediatrics (AAP), European Society of Neuroradiology (ESNR), American Professional Society on the Abuse of Children (APSAC), Swedish Paediatric Society, Norwegian Pediatric Association and Japanese Pediatric Society addresses significant misconceptions about the diagnosis of abusive head trauma (AHT) in infants and children. It builds on 15 major national and international professional medical societies’ and organizations’ consensus statements confirming the validity of the AHT diagnosis. The statement also exposes the fallacy of simplifying the diagnostic process to a “triad of findings” — a legal argument and not a medically valid term.

AHT is the leading cause of fatal head injuries in children younger than 2 years and is responsible for 53% of serious or fatal traumatic brain injury cases. The etiology of injury is multifactorial (shaking, shaking and impact, impact, etc.) so that the current best and most inclusive term is AHT, as advanced by the American Academy of Pediatrics.

No single injury is diagnostic of AHT. Rather the multiplicity of findings including evidence of intracranial and spinal involvement, complex retinal hemorrhages, rib and other fractures inconsistent with the provided mechanism of trauma, as well as the severity and age of the findings provide clues to the diagnosis. Subdural hematoma is the most frequently identified intracranial lesion but brain parenchymal injury is the most significant cause of morbidity and mortality in this setting. There is a high incidence of ligamentous cervical spine injury among victims of inflicted injury. However, it is important to emphasize that absence of ligamentous injury does not exclude AHT. In suspected cases of AHT, alternative diagnoses must be considered and when appropriate explored. The question to be answered is, “Is there a medical cause to explain all the findings or did this child suffer from inflicted injury?”

Despite courtroom arguments by defense lawyers and their retained physician witnesses, there is no reliable medical evidence that the following processes are precise mimics or causative in the constellation of injuries characteristic of AHT: cerebral sinovenous thrombosis, hypoxic–ischemic injury, lumbar puncture or dysphagic choking/vomiting. There is also no substantiation, at a time remote from birth, of the proposal that birth-related subdural hemorrhages can result in sudden collapse, coma or death caused by acute rebleeding into a previously asymptomatic chronic collection. In addition, subdural hematoma is uncommon in the setting of benign enlargement of the subarachnoid space (BESS), and when subdural hematoma is present, AHT should be considered.

The diagnosis of AHT is a medical diagnosis made by a multidisciplinary team of pediatricians and pediatric subspecialty physicians, social workers and other professionals based on consideration of all the facts and evidence. AHT is a scientifically non-controversial medical diagnosis broadly recognized and managed throughout the world. When diagnosed, it signifies that accidental and disease processes cannot plausibly explain the etiology of the infant/child’s injuries. A diagnosis of AHT is a medical conclusion, not a legal determination of the intent of the perpetrator or, in the false hyperbole of the courtroom and sensationalistic media, “a diagnosis of murder.”

The question in civil and criminal court cases involving allegations of unwitnessed abuse is the quality of the medical evidence and the integrity and expertise of the medical witness’s testimony. Over the last decade, the courtroom has become a forum for medical opinions on the etiology of infant/child head injuries that runs the gamut from the well-founded evidence-based conclusions of multidisciplinary medical teams to speculative theories that cannot be reconciled with the medical evidence that is generally accepted in the relevant medical community. When pivotal medical testimony is contradictory, the message to the courts, the news media and the general public about infant injuries and safe caregiving is often confusing and inaccurate.

Professional medical societies use consensus statements to communicate general physician acceptance on a particular topic. These statements are vetted by the membership and designed to help physicians, news media and the public distinguish accurate medical information from non-evidence-based or “courtroom-only” causation theories. The formal
This consensus statement reviews and synthesizes relevant scientific data. This statement is supported by the SPR Child Abuse Imaging Committee and endorsed by the boards of directors of the Society for Pediatric Radiology (SPR), European Society of Paediatric Radiology (ESPR), American Society of Pediatric Neuroradiology (ASPNR), American Academy of Pediatrics (AAP), European Society of Neuroradiology (ESNR), American Professional Society on the Abuse of Children (APASAC), Swedish Paediatric Society, Norwegian Pediatric Association and Japanese Pediatric Society. This statement is derived from an empirical assessment of the quality and accuracy of the medical literature and addresses the threshold question of when such literature is generally medically accepted in the pediatric health care community. This review of the medical literature also considers the court admissibility and the reliability of expert medical opinions based on such literature. The contributing board-certified physician authors each have one or more pediatric subspecialty board certifications from the American Board of Radiology or the American Board of Pediatrics or American Board of Neurosurgery (all member organizations of the American Board of Medical Specialties) or Royal College of Radiologists (UK) or equivalent boards in Greece and Italy. Additionally, all authors have 10–40 years of individual clinical experience diagnosing and treating children. The non-physician author is a law professor with nearly two decades of experience researching and writing on the appropriate use of child abuse evidence in court.

We address the following questions:

1. What are the causes of head injury in infants and young children? Why has AHT terminology evolved (shaken baby syndrome, battered child, abusive head trauma, etc.)?
2. What are the presenting features of AHT?
3. How is the diagnosis of AHT made?
4. What unsubstantiated alternative diagnoses are being proffered in the court?
5. What is the role of the multidisciplinary child protection team in the determination of AHT?
6. What are the issues that perpetuate misconceptions in the courtroom?
7. What can be done to provide the courts accurate information about the state of medical knowledge in AHT?

Etiology of head trauma in infants and young children and nomenclature of abusive head trauma (AHT)

When data are evaluated from head trauma in children younger than 2 years old, AHT is recognized as the leading cause of fatal head injuries and is responsible for 53% of the serious or fatal traumatic brain injury cases [8]. The peak incidence of

Introduction

This consensus statement addresses significant misconceptions and misrepresentations about the diagnosis of abusive head trauma (AHT) in infants and young children. Major national and international professional medical societies and organizations have consistently confirmed the validity of the AHT diagnosis, its classic features and its severity [1–4].

Recently, denialism of child abuse has become a significant medical, legal and public health problem. In courtrooms in the United States defense attorneys and the medical witnesses who testify for them have been disseminating inaccurate and dangerous messages that are often repeated by the news media. Instead of arguing that there is reasonable doubt that physicians made a mistake in this case, they are arguing that child abuse is routinely overdiagnosed. The deliberate dissemination of this misinformation will deter caregivers from seeking medical services for infants and children — even in cases where there has been no abuse or neglect. The accompanying defense message — that shaking an infant cannot cause serious injury — will create the additional risk of encouraging dangerous or even life-threatening caregiver behavior. The majority of the expert witnesses practice evidence-based medicine; they base their testimony on clinical expertise and peer-reviewed evidence in the medical literature. However in some legal AHT cases, defense arguments (frequently supported by opinion testimony provided by a small group of medical witnesses) have offered a scientific-sounding critique of the AHT diagnosis by offering a laundry list of alternative causation hypotheses [5]. Efforts to create doubt about AHT include the deliberate mischaracterization and replacement of the complex and multifaceted diagnostic process by a near-mechanical determination based on the “triad” — the findings of subdural hemorrhage, retinal hemorrhage and encephalopathy [1]. This critique has been sensationalized in the mass media in an attempt to create the appearance of a “medical controversy” where there is none [6, 7]. The straw man “triad” argument ignores the fact that the AHT diagnosis typically is made only after careful consideration of all historical, clinical and laboratory findings as well as radiologic investigations by the collaboration of a multidisciplinary team.

Dissemination of this information via a consensus statement is intended to help courts improve the scientific accuracy of their decisions involving vital public health issues. Consensus statements reduce confusion by recommending to judges and jurors the tools necessary to distinguish genuine evidence-based opinions of the relevant medical community from legal arguments or etiological speculations that are unwarranted by the clinical findings, medical evidence and evidence-based literature.
fatal AHT is at 1–2 months of age [9]. Terms used to describe this form of head injury have evolved as scientific data have advanced [10] (Table 1 with references [11–16]). This abusive form of head trauma occurs most frequently with other forms of abuse and less often in isolation [17].


In 1972 and again in 1974, Caffey [14, 15] postulated that the practice of “whiplash shaking and jerking of abused infants are common causes of the skeletal as well as the cerebrovascular lesion.” He referred to the earlier work of Ommaya and Yarmell [22] and that of Guthkeleh [23] to show the effects of rotational acceleration/deceleration of whiplash as the etiology of subdural hematomas. This mechanism explains why there are frequently no external marks of injury and also provides a reason for the retinal hemmorhages found in abused children [24–26]. In these papers, Caffey mentioned that whiplash/shaking may cause “protracted, repeated breath holding spells which may be similarly damaging to the brain” and was prescient to theories and data published decades later regarding hypoxic–ischemic injury associated with AHT [14, 15, 27–29]. Of note, whiplash/shaking has been repeatedly reaffirmed by confessions of perpetrators in which violent shaking was the most commonly reported mechanism of injury (68–100%) [30–32].

In 1987, Duhaime et al. [16] postulated that based on clinical, pathological data and biomechanical models, rotational acceleration/deceleration whiplash injuries do not provide enough force to account for the severe injuries of these children and that in severe cases blunt trauma must be involved. From this article, the term shaken baby/shaken impact emerged. There still remains discussion over whether shaking alone or shaking with blunt trauma is necessary for the injuries of these abused children, but confessional evidence is quite striking that shaking alone can cause AHT [30–32]. Dias [33] made the case that shaking alone can be a causative mechanism and significantly questioned the validity of the biomechanical model of Duhaime et al. [16]. In 2016, Narang et al. [3] documented that both AHT and shaken baby syndrome (SBS) are generally accepted diagnoses in the medical community. Currently, the medical literature and overwhelming clinical experience and judgment demonstrate that AHT can be caused by shaking alone, shaking with impact, or blunt impact alone.

In 2009, the Committee on Child Abuse and Neglect of the American Academy of Pediatrics issued a statement recommending the medical use of the term abusive head trauma (AHT) [10]. This policy statement did not negate the mechanism of shaking as a significant mechanism of injury but instead merely clarified that the term “shaking” alone was not inclusive of the full range of injury mechanisms. AHT is the most comprehensive term for the intracranial and spinal lesions in abused infants and children. In various forms, AHT has been in the modern medical literature for more than 60 years [34], “with over 1,000 peer-reviewed clinical medical articles written by over 1,000 medical authors from more than 25 different countries” [2]. Inflicted brain injuries are multifactorial in origin. It is the role of physicians to determine whether the injuries and the history for the injuries are suspicious for AHT and whether the child should be evaluated by a multidisciplinary child protection team with the goal of protecting the child. We note that the repeated defense counsel argument that the 2009 AAP statement constitutes a rejection of the medical evidence for shaking as a mechanism of infant injury is false and misleading legal rhetoric without any factual support in the statement or in any other statement from the AAP.

**The presenting features of AHT**

The clinical presenting features of AHT include severe head injury; death; less severe trauma with an unexplained mechanism; unsuspected finding on imaging or assessment for macrocephaly, developmental delay, seizures or other neurologic concerns; or discovery during the workup as a sibling of an abused child. The clinical findings might include neurologic signs and symptoms such as irritability/lethargy, altered mental status, seizures, respiratory compromise and apnea,
fractures, varying degrees of pattern marks or bruises in unusual locations, vomiting and poor feeding [35].

Children with fatal head injuries have altered mental status immediately after the injury [36]. However on rare occasions young victims of fatal head trauma present with Glasgow coma scale (GCS) of >12 for a short time before death, although GCS is a very rough guide of normalcy in the youngest age group [36, 37]. There is no evidence that children with fatal head trauma have prolonged asymmetric lucid intervals prior to neurologic collapse. Some victims of AHT who sustain non-fatal injuries have nonspecific symptoms for several hours or more before developing either seizures or coma, while others remain relatively asymptomatic. Sixty-five percent of AHT cases present with neurologic abnormality while the remainder present with nonspecific symptoms [38]. This lack of specificity and other factors can lead to inaccurate diagnosis unless the evaluating physician understands the broad clinical spectrum of AHT [39].

Kemp et al. [40] described the predictive power of different neuroradiologic features to aid in the distinction of AHT from other causes. The clinical certainty for AHT is higher for children with more severe presentations or with multiple findings [17, 41]. Several characteristic findings have most frequently been identified in AHT including subdural hematoma (SDH), brain parenchymal injuries, retinal hemorrhages and rib fractures [2, 10, 41, 42]. In the review by Maguire et al. [41], any combination of three or more of the significant diagnostic features yielded a positive predictive value of 85%. Kelly et al. [43], in their review of referrals to a child protection team over a 20-year period, reported that in children younger than 2 years the characteristics of particular interest for AHT included no history of trauma (90%), no external evidence of impact to the head (90%), complex skull fractures with intracranial injury (79%), subdural hemorrhage (89%) and hypoxic–ischemic injury (97%).

How the diagnosis of AHT is made

The diagnosis of AHT is made like any other medical diagnosis, by considering all the information acquired via clinical history, physical examination, and laboratory and imaging data.

History

Inconsistency of the presenting history with the clinical findings is a concern for child maltreatment including AHT. Therefore, detailed history including a follow-up history once the acute illness has been addressed is vital to diagnostic accuracy [44, 45]. The two most common histories provided in cases of confirmed AHT are a low-height fall (of less than 4–6 ft) and no specific history of trauma [46]. Severe head injury or moderate to large non-focal SDH are rarely consistent with a history of a short fall of less than 4 ft [47].

There are significant limitations with published biomechanical studies evaluating falls including a lack of complete biofidelic integrity [48–51]. The data for injury thresholds in these studies were derived from adult primates undergoing single, non-impact accelerations [48–51]. The differences in intrinsic material properties of the infant skull, brain, cerebrospinal fluid (CSF) and blood vessels versus an adult human or primate were not considered, nor were the effects of repeated injury [33]. We need to develop a better understanding of these critical differences to develop better biomechanical studies approximating real-life situations in order to provide more accurate and reliable information.

Review of extensive literature demonstrates that severe intracranial injury from short falls is rare, and the predictions from any biomechanical study/model should not deviate too much from established extensive real-life data to be considered valid [25, 47, 52–86]. For example, Chadwick et al. [52] in their study of short falls demonstrated a mortality of 0.48 per million per year in children younger than 5 years. A review of 26 studies of accidental falls from various heights [25, 72–85] involving 1,902 children found 23 fatal injuries, of which only 0.26% (5/1,902) were from falls less than three stories [47]. In a review of 24 in-hospital newborn falls from less than 1 m height, 2 babies had non-depressed linear parietal fractures and 2 babies without skull fracture had infratentorial SDH, which was thought to be birth-trauma-related SDH and unrelated to the fall. All the babies had a normal or benign physical examination post fall and had normal findings on examination at discharge [86].

Review of the extensive literature informs us that mortality from short falls is extremely rare, and the majority of these are benign occurrences with no significant neurologic dysfunction. Linear skull fracture, associated epidural hemorrhage, focal contusion and rarely small focal SDH or subarachnoid hemorrhage might be seen on imaging, but significant intracranial hemorrhage, parenchymal contusion or diffuse hypoxic–ischemic injury is uncommon in contrast to findings seen in AHT. When significant neurologic dysfunction or mortality does occur with short falls, it is related to a large extra-axial hematoma or vascular dissection and secondary stroke [33, 52].

Physical examination and importance of ocular findings

Clinicians should perform a meticulous examination for external bruises and tenderness. Bruises to the head and face have been associated with AHT, and patterns of injury consistent with grabbing, choking and blunt trauma should be sought [69, 87]. The absence of external trauma to the head and neck is common, however, and sometimes
soft-tissue injuries including scalp hematomas are only evident at autopsy [88].

Ocular findings in AHT include orbital and lid ecchymosis, subconjunctival hemorrhage, anisocoria and disconjugate eye movements and retinal hemorrhages. Retinal hemorrhages are an important finding in AHT and when abuse is suspected, a prompt complete examination including full indirect ophthalmoscopic examination through a dilated pupil should be obtained [87]. The incidence of retinal hemorrhage in AHT is approximately 85% [89, 90]. “Hemorrhages that are too numerous to count, multilayered and extending to the ora serrata are specific” [91]. A number of conditions have been associated with retinal hemorrhages, but this quoted description is highly suspicious for AHT [87] (Table 2; also see reference [92]). The retina is multilayered and traumatic retinoschisis occurs from vitreo-retinal traction sustained from repeated rapid acceleration/deceleration forces [93]. Deep splits of the retina and even focal retinal detachment can occur. Retinal folds are hypopigmented ridges, usually around the macula. In the absence of severe documented head trauma, retinal folds and retinoschisis are more specific for AHT [93]. These types of retinal lesions do not occur from birth trauma or papilledema (papilledema occurs in 10% of AHT) [87].

A prompt evaluation for retinal hemorrhages is important because they can fade rapidly. Generally, intraretinal hemorrhages clear rapidly, whereas preretinal hemorrhages might persist for many weeks [94]. The presence of too-numerous-to-count intraretinal hemorrhages might indicate that trauma occurred within a few days prior to examination, whereas the presence of preretinal with no or few intraretinal hemorrhages suggests days to weeks since trauma [94]. To identify these patterns accurately, the health care team should complete eye examinations as soon as possible after admission, preferably within 24–48 h [94].

Laboratory studies and imaging

Although the history and physical examination are paramount, appropriate use of laboratory studies and imaging is vital for accurate diagnosis and treatment. Recent papers discuss the evaluation of bleeding and bone diseases when there is a suspicion of abuse [95, 96]. Skeletal survey following current guidelines should be performed for all children with potential AHT, particularly those younger than 2 years [4]. In older children, long-bone fractures can be more reliably suspected in the presence of extremity tenderness, swelling or refusal to bear weight.

For an acutely ill child with neurologic impairment, an optimal imaging strategy involves initial unenhanced CT with 3-D reformatted images of the calvarium [97], followed by a full multi-sequence MRI of the brain and the cervical, thoracic and lumbar spine as soon as feasible. Children who are intact neurologically can be imaged with MR first [98–101]. Suspicion of AHT warrants comprehensive imaging, and the decision rule developed from a network of emergency departments regarding the use of imaging in low-risk blunt head trauma does not apply when there are concerns for AHT [102–104]. Intracranial bleeding is common in AHT and often presents as subdural hematoma. Magnetic resonance imaging of the brain and spine with a variety of sequences is useful in characterizing extra-axial bleeds and defining cerebral contusion, laceration and other parenchymal brain injuries.

A number of comparative studies in young children have elucidated the statistical differences in the types and severity of intracranial injuries from accidental versus abusive head trauma [25, 32, 43, 46, 72, 76, 77, 79, 83, 105–108]. These studies collectively demonstrate that: (1) skull fractures are equally common following accidental trauma and AHT, but the complex skull fractures are more common following AHT; (2) epidural hematomas are more common following accidental trauma; (3) subdural hematomas are far more common following AHT; and (4) subarachnoid, intraparenchymal and intraventricular hemorrhage are equally common in both AHT and accidental trauma [25, 32, 43, 46, 72, 76, 77, 79, 83, 105–107].

Subdural hematoma is the most commonly observed intracranial lesion (in up to 90%) in young infants with AHT and is most commonly parafalcine in location [109, 110]. The inflicted injury (acceleration/deceleration +/- impact) can lead to tearing of convexity bridging veins at the junction of the bridging vein and superior sagittal sinus. Additionally, rupture of the arachnoid membrane allows cerebrospinal fluid to enter the subdural space, mixing with subdural blood (hematohygroma) [111, 112]. SDH might have a mixed attenuation at presentation (Table 3). Mixed-attenuation subdural hematomas are found with greater prevalence in AHT than in accidental head trauma [109]. In a review by Bradford et al. [110], of 105 confirmed AHT cases, intracranial SDH was identified in 92% of cases. On the initial diagnostic CT study, the SDH was of homogeneous hyperattenuation in 28% of cases, mixed attenuation in 58% of cases and homogeneous hypoattenuation in 14% of cases. In the cases with homogeneous hyperattenuation SDH on the initial CT, the first hypoattenuated component was seen between 0.3 days and 16 days after injury and the disappearance of the last hyperattenuated component was identified between 2 days and 40 days after injury. For these reasons, precise estimation of age of the mixed-attenuation SDH on the initial CT should be avoided.

While SDH is the most frequent intracranial lesion in AHT, parenchymal brain injury is the most significant cause of morbidity and mortality [113]. The injury might be direct mechanical injury such as contusion, direct axonal injury, laceration or parenchymal hematoma or indirect in nature, resulting from hypoxia and ischemia [113]. MRI is more sensitive than CT in delineation of parenchymal injuries. Timing parenchymal and

---

[References]

72, 76, 77, 79, 83, 105–107
70
111, 112
98–101
105–108
72, 76, 77, 79, 83
109, 110
102–104
95, 96
[95, 96]
[97]
[98–101]
extra-axial injury can be challenging, and because injuries evolve over time, repeat MRI is frequently indicated.

Venous injury is strongly associated with AHT. It is common at the junction of the bridging vein and superior sagittal sinus complex and is considered the source of SDH [109, 114]. Choudhary et al. [114] found that nearly 70% of children with AHT had some sort of venous abnormality. Findings consisted of cortical vein injury (44%) and mass effect on cortical draining veins or dural sinuses (69%). Specifically, disruption of bridging veins at their insertion into the superior sagittal sinus is a common source of SDH in AHT. Rupture of smaller intradural vessels resulting in subdural hemorrhage, likely caused by trauma, has also been proposed as an etiology [115, 116]. Trauma of both types, accidental and AHT, causes venous injury including intracranial venous thrombosis.

Young infants are at an increased risk of upper cervical spinal injury. Such injury is more likely to be soft-tissue or ligamentous in nature [117]. Imaging of bony cervical spine is infrequently positive (0.3–2.7%) in children investigated for suspected child abuse [118]. Non-bony spinal abnormalities have, however, been identified in up to 2/3 of victims of AHT, in both clinical and autopsy series [117, 119, 120]. Choudhary et al. [119] has shown on MRI that 78% of these infants have spinal findings, mostly ligamentous, and up to 75% have spinal subdural hematoma that tracks from the posterior fossa [117, 119, 121]. It is apparent that cervical, thoracic and lumbar MRI should be included in the diagnostic workup when there is evidence of intracranial injury. Prior to knowledge of the ligamentous injury, those who denied the existence

### Table 2

<table>
<thead>
<tr>
<th>Injury or condition</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accidental trauma</td>
<td>Few in number except in very severe trauma, usually limited to posterior pole, predominantly intraretinal and pre-retinal, extremely rare (most studies &lt;3% incidence) after short falls except if there has been an epidural hemorrhage or occipital impact</td>
</tr>
<tr>
<td>Birth</td>
<td>Between 19.2% and 37.3% incidence in vaginal birth, 6% incidence after C-section</td>
</tr>
<tr>
<td>Motor vehicle crash or severe crush injury</td>
<td>Easily determined by history</td>
</tr>
<tr>
<td>Cardiopulmonary resuscitation</td>
<td>Extremely rare, few in number, posterior pole</td>
</tr>
<tr>
<td>Extracorporeal membrane oxygenation (ECMO)</td>
<td>5 of 37 (13%) ECMO patients had retinal hemorrhage</td>
</tr>
<tr>
<td>Prematurity</td>
<td>Retinal hemorrhage occurs at the peripheral circumferential demarcation between the vascularized and avascular retina</td>
</tr>
<tr>
<td>Intracranial hypertension or papilledema</td>
<td>Small number of retinal hemorrhages on or around the optic disc</td>
</tr>
<tr>
<td>Coagulopathy/anemia</td>
<td>Uncommon, few in number, posterior pole severe anemia and usually thrombocytopenia required, often with cotton wool spots*</td>
</tr>
<tr>
<td>Meningitis</td>
<td>More often if coagulopathy or sepsis is present. Only severe retinal hemorrhage if purulent meningitis, otherwise few in number, posterior pole</td>
</tr>
<tr>
<td>Ruptures aneurysm/arteriovenous malformation</td>
<td>May have severe extensive retinal hemorrhage; vascular malformation easily recognized on neuroimaging</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>Few in posterior pole</td>
</tr>
<tr>
<td>Menkes disease</td>
<td>Causes blue sclera</td>
</tr>
<tr>
<td>Galactosemia</td>
<td>Vitreous hemorrages reported</td>
</tr>
<tr>
<td>Glutaric aciduria</td>
<td>Rarely occurs and is confined to posterior pole</td>
</tr>
</tbody>
</table>

* Rare in critically ill children with fatal accidental trauma, severe coagulopathy sepsis and myeloid leukemia [92]

### Table 3

<table>
<thead>
<tr>
<th>Appearance of subdural collection on CT</th>
<th>Possible time frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iso-attenuation</td>
<td>Hyperacute, acute</td>
</tr>
<tr>
<td>Hyperattenuation</td>
<td>Acute, early subacute</td>
</tr>
<tr>
<td>Mixed hyper- and hypoattenuation</td>
<td>Hyperacute, acute, subacute and chronic</td>
</tr>
<tr>
<td>Hypoattenuation</td>
<td>Chronic</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Processes associated with retinal bleeding (modified from Levin et al. [87])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injury or condition</td>
</tr>
<tr>
<td>Accidental trauma</td>
</tr>
<tr>
<td>Birth</td>
</tr>
<tr>
<td>Motor vehicle crash or severe crush injury</td>
</tr>
<tr>
<td>Cardiopulmonary resuscitation</td>
</tr>
<tr>
<td>Extracorporeal membrane oxygenation (ECMO)</td>
</tr>
<tr>
<td>Prematurity</td>
</tr>
<tr>
<td>Intracranial hypertension or papilledema</td>
</tr>
<tr>
<td>Coagulopathy/anemia</td>
</tr>
<tr>
<td>Meningitis</td>
</tr>
<tr>
<td>Ruptures aneurysm/arteriovenous malformation</td>
</tr>
<tr>
<td>Hypoxia</td>
</tr>
<tr>
<td>Menkes disease</td>
</tr>
<tr>
<td>Galactosemia</td>
</tr>
<tr>
<td>Glutaric aciduria</td>
</tr>
</tbody>
</table>

* Rare in critically ill children with fatal accidental trauma, severe coagulopathy sepsis and myeloid leukemia [92]
of the shaken baby mechanism used “lack of spinal injury” to boost their unfounded theory [122–124]. However, it is important to emphasize that absence of ligamentous injury does not exclude AHT.

Unsubstantiated alternative theories proffered in the court [109]

The determination of whether certain theories are putative explanations for AHT must at least recognize the long and storied medical history of the many etiologies already investigated as reasonable explanations. With those historical investigations as a foundation, trauma has come to be uniformly recognized as the primary etiology of pediatric and adult SDHs [46]. Depending on the health history, clinical presentation and pertinent laboratory testing, there are diseases that are considered in the differential diagnosis of subdural hematoma and appropriate medical evaluation is required for all children.

Because medicine and science are dynamic, it is important to continually evaluate new hypotheses and, consequently, re-evaluate previously confirmed scientific understanding, thus avoiding a rush to judgment. In this section, we discuss selected current theories proffered as causative bases for AHT that reportedly “mimic” the injuries seen. However, the lack of scientific evidence for these assertions underscores the general consensus opinion of pediatricians and pediatric subspecialists against these theories as reasonable explanations for AHT [1, 125]. Most of these unsubstantiated alternative theories just focus on one aspect of the range of injuries seen in AHT while conveniently ignoring other injuries that cannot be explained away. For instance, those postulating cerebral sinovenous thrombosis (CSVT) theory as an alternative diagnosis of AHT focus on retinal hemorrhage and intracranial SDH while they ignore concomitant skeletal injuries, neck injury and visceral injury.

The theories have included association of common procedures such as lumbar puncture and common symptoms such as cough with uncommon clinical presentations such as CSVT or hypoxic–ischemic injuries (HII) in the newborn. The theory of lumbar puncture leading to intracranial hemorrhage precisely mimicking AHT speculates that loss of CSF pressure leads to intracranial hypotension and resultant SDH, but the only evidence provided has been couple of case reports in older children and adult literature [126–128]. Meanwhile lumbar puncture is a routine procedure performed safely across outpatient and inpatient settings without intracranial sequelae. Complications from lumbar puncture are rare, and in fact a recent study in adults has documented that an underlying issue such as coagulopathy is typically present when complications arise [129].

Similarly, sustained cough, choking and dysphagic choking have been speculated to cause SDH and retinal hemorrhage mimicking AHT. The theory speculates that any cause of sustained raised intrathoracic pressure such as choking, paroxysmal coughing, gagging or vomiting can cause increased intracranial and retinal venous pressure by impeding thoracic venous return, leading to traumatic venous rupture with retinal hemorrhage and SDH [130, 131]. However a computer model developed to prove this hypothesis was lacking because it did not have a clearly defined threshold for failure of bridging vein in infants and because it was developed from data obtained mostly from adult and animal studies [109, 131]. An isolated case report of SDH in an infant with pertussis has also been cited to support this theory, but this particular case also had a confounding history of a fall a week before presentation, which might have been responsible for the SDH [109, 132]. Additionally, this theory has been negated by prospective studies in 83 infants suffering from pertussis demonstrating no evidence of retinal hemorrhages seen in AHT [133, 134]. Dysphagic choking-type of acute life-threatening event (ALTE) mimicking AHT was described in a Barnes et al. [135] case report and review [136]. The case report has been criticized for failing to disclose the source of information, for the author’s role as defense expert witness, for omission and misrepresentations of certain facts and legal outcome, for lacking proper evidence base and for use of inaccurate information to support speculative explanations [137, 138]. ALTE, which has been replaced with the new terminology “brief resolved unexplained events,” has been shown to have a low prevalence of retinal hemorrhage or SDH and cannot be considered to be the cause of SDH or retinal hemorrhage [139–141]. Similarly, retinal hemorrhage was not identified in a prospective study of vomiting infants with hypertrophic pyloric stenosis [142]. These prospective studies underline the fact that while the cough/dyphagic choking/vomiting theory is supported by no recent solid evidence base, there are strong prospective studies providing evidence that refutes these theories. In a retrospective study, children who presented with ALTE and subdural hemorrhages were found to be nearly 5 times more likely to have at least one suspicious extracranial injury, supporting the diagnosis of AHT and thereby negating the role of ALTE as a causative mechanism for findings concerning AHT [141].

Hypoxic–ischemic injury is another diagnosis proposed as an etiology of intracranial SDH and retinal hemorrhage, posited by some to precisely mimic AHT [143, 144]. This is based upon Geddes et al.’s [143] unified hypoxia theory, which derived its findings from the commonality between intracranial postmortem findings of pediatric patients who suffered from hypoxia and people with AHT. However, this theory has been refuted by a number of studies in which SDH was not identified on pathology or imaging or either in the clinical context of hypoxic injury [145–148]. Besides, traumatic AHT can be present without hypoxia, and AHT with hypoxic injury can coexist with other clinical findings such as visceral or skeletal injuries.
and paraspinal soft-tissue injuries supporting the diagnosis of AHT [117]. Although hypoxia is frequently seen in traumatic injury of the brain, it is likely a comorbid association similar to other traumatic injuries of the brain and spine.

Cerebral sinovenous thrombosis has been proposed as a cause of intracranial injury in children. This unsupported theory proposes that raised intracranial venous pressure resulting from cerebral sinovenous thrombosis leads to bursting of bridging veins resulting in brain parenchymal injury, SDH and retinal hemorrhage similar to the pattern of injuries seen in AHT [114, 149–151]. CSVT is an uncommon disorder in childhood but fortunately has been well reported in the literature and thereby provides us with a robust evidence base to conclusively refute this theory [109, 152–157]. Although CSVT has been associated with parenchymal hemorrhagic infarct, resulting in significant morbidity and mortality, there is no evidence in the literature where primary CSVT thrombosis has been identified as the cause of acute SDH or a presentation with abrupt collapse with prolonged coma in a previously healthy child [114]. CSVT has been identified in situations where it is secondary in nature, consistent with the mechanism of pathology such as iron deficiency anemia or an inherited predisposition toward coagulation and trauma [109, 114]. We should not confuse thrombosis with subcortical hemorrhage; similarly, absence of veins on MR venogram doesn’t equate to thrombosis, and demonstration of intraluminal thrombosis is equally important [114].

Subdural hematoma in the setting of benign enlargement of the subarachnoid space (BESS)

Benign enlargement of subarachnoid space (BESS) is common in the setting of macrocephaly in infancy. Although BESS was initially thought to predispose children to SDH with minimal trauma [158], the latest reviews reveal that less than 6% of infants with BESS develop hemorrhagic subdural collections (Table 4, references [158–164]). Most of the published series are lacking because of their variable methods of ascertainment, variable descriptions of the kind of subdural collections — cerebrospinal fluid, hemorrhagic fluid or a mixture of the two — and incomplete assessment for abuse in these cases [162].

Taking only those reports from Table 4, in which the prevalence of BESS has also been documented, a total of 712 cases of BESS were documented, with 38/712 (5.3%) reported to have subdural collection, including 12/712 (1.7%) that were reported to be hemorrhagic in nature. Accidental trauma or abuse was reported in 5/12 (41.7%) of the subdural collections that were hemorrhagic. Besides, up to 50% of children with BESS and SDH may display concomitant important injuries [165]. Overall subdural collections are uncommonly seen in the setting of BESS and assessment to exclude trauma, including AHT, should be performed in those with hemorrhagic and non-hemorrhagic subdural collections, especially in children younger than 2 years.

Birth trauma

The risk factor for intracranial hemorrhage in newborn infants is abnormal labor, as evidenced by a higher rate of traumatic brain injury in infants born by Cesarean section after an abnormal labor and those born with vacuum extraction and forceps as compared to infants born by spontaneous vaginal delivery or delivered by elective Cesarean section [166]. Birth trauma accounts for 1–2% of the mortality in newborns and any significant intracranial injury presents in the immediate postnatal period with significant clinical symptoms such as

<table>
<thead>
<tr>
<th>Authors</th>
<th>Number of patients with BESS</th>
<th>Number of subdural collections (% of total BESS cases)</th>
<th>Number (% of total BESS) with reported hemorrhagic subdural collections</th>
<th>Other details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilms et al. [158]</td>
<td>19</td>
<td>8 (42.1%)</td>
<td>3 (15.8%)</td>
<td>One case of recent trauma with hemorrhagic subdural collection</td>
</tr>
<tr>
<td>McKeag et al. [159]</td>
<td>177</td>
<td>4 (2.3%)</td>
<td>4 (2.3%)</td>
<td>1 rib fracture</td>
</tr>
<tr>
<td>Tucker et al. [160]</td>
<td>311</td>
<td>18 (5.8%)</td>
<td>1 (0.3%)</td>
<td>Hemorrhagic subdural collection case reported for abuse</td>
</tr>
<tr>
<td>Greiner et al. [161]</td>
<td>108</td>
<td>6 (5.6%)</td>
<td>2 (1.9%)</td>
<td>2 reported for abuse</td>
</tr>
<tr>
<td>Mcneely et al. [162]</td>
<td>n/a</td>
<td>7 (n/a)</td>
<td>7 (n/a)</td>
<td>Abuse cases were excluded. 2 cases with accidental trauma</td>
</tr>
<tr>
<td>Haws et al. [163]</td>
<td>84</td>
<td>2 (2.4%)</td>
<td>2 (2.4%)</td>
<td>n/a</td>
</tr>
<tr>
<td>Alper et al. [164]</td>
<td>13</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

(n/a=not available)
irritability, poor feeding, emesis, apnea or disordered breathing, bradycardia, and seizures or disordered mentation [167–184].

Small birth-related subdural hematomas, most commonly along the tentorium, parietal occipital convexity, retrocerebellar posterior fossa or interhemispheric fissure, are observed in 8–46% of asymptomatic newborn infants [185–187]. This has led to the unsubstantiated theory that rebleeding, months later, in persistent birth-related asymptomatic SDH can present acutely with clinical features mimicking AHT [188]. Rooks et al. [186] in 2008 reported MRI findings within 72 h of birth and serial developmental evaluations of 101 asymptomatic neonates, 79 born by vaginal delivery and 22 by Cesarean delivery. SDH was present in 46 (46%) of the infants, most of whom resolved on follow-up MRI by 1 month and all resolved by 3 months. There were no significant differences in clinical outcomes in this cohort, as compared to the normal population, on serial developmental examinations [186]. Other authors have reported similar findings [187, 189].

To summarize, asymptomatic birth-related subdural hematomas are relatively frequent and resolve in the overwhelming majority of infants within the first 4–6 postnatal weeks, and do not appear to rebleed. If there is significant birth-related trauma, neonates are symptomatic in the immediate postnatal period. In particular, there is no merit to the unsubstantiated proposal that acute collapse, coma or death, occurring months after delivery, is caused by a parturitional SDH with secondary rebleeding.

Multidisciplinary assessment and long-term outcome

The medical diagnosis of AHT is made by pediatricians and pediatric subspecialists based on medical evaluation. In many children’s hospitals, an interdisciplinary team of specialists that includes physicians, nurses, hospital social workers and others works together to evaluate cases. Hospital-based multidisciplinary teams have existed in many communities to provide comprehensive assessments and services for families for more than 60 years. The overriding goal of the work of these teams is to diagnose and to treat child abuse and neglect, assess for alternative diagnoses when appropriate, and assist in the efforts of the many agencies involved. The Children’s Hospital Association (formerly the National Association of Children’s Hospitals and Related Institutions) has released guidelines for team composition and function to aid in providing services [101, 190]. In addition, in some jurisdictions, multidisciplinary teams of hospital and community professionals review injuries, medical history, family and social risk to reach a more comprehensive assessment. These hospital–community partnerships are composed of physicians, nurses, social workers, clergy, psychologists, child protection services, law enforcement and other professionals with relevant experience. These multidisciplinary teams can review all of the data related to the case from different perspectives to gain a more complete understanding of the issues [8, 45, 191–194]. Whenever members of these teams present testimony in a legal setting, there has usually been much in-depth consideration of the diagnosis, and the probability of the correct diagnosis is high.

Abusive head trauma is the leading cause of physical abuse fatalities. In a review of child abuse fatalities, the authors identified shaking as a cause or contributor in 45% of the deaths, with beating, kicking and chronic battering accounting for the rest [191]. The authors identified crying as the trigger for 20% of deaths, followed by disobedience (6%), domestic arguments (5%), toilet training (4%) and feeding problems (3%) [191]. Infants are significantly more likely to be physically abused when a caretaker has an emotional disturbance and when there is violence between caretakers [195]. Unfortunately, when AHT is not prevented, the outcome can be devastating (Table 5) and the financial costs to society extremely high [196]. The estimated lifetime cost of 4,824 cases in 2010 was $13.5 billion [197].

The issues that perpetuate misconceptions in the courtroom

The most recent AAP policy statement on expert witness testimony has reemphasized the fact that expert witness neutrality and professional integrity can be pivotal factors in civil and criminal child abuse cases [198]. When expert testimony is scientifically reliable, objective and accurate, it provides useful information for the legal factfinder. Ethical and professional norms of responsible expert testimony require that physicians be objective and neutral assessors and conveyors of medical information, which means that they weigh the scientific merit of their opinions and conclusions and “present testimony that reflects the generally accepted standard within the specialty or area of practice, including those standards held by a significant minority” [198, 199]. Regrettably, not all medical experts’ courtroom testimony falls within these ethical and professional boundaries. A few physicians, including those who do not treat or diagnose children as part of their medical

Table 5  Outcomes after abusive head trauma [196]

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death (20-25%)</td>
<td></td>
</tr>
<tr>
<td>Spastic hemiplegia or quadriplegia</td>
<td>(15–64%)</td>
</tr>
<tr>
<td>Intractable epilepsy (11–32%)</td>
<td></td>
</tr>
<tr>
<td>Microcephaly with cortico-subcortical atrophy (61–100%)</td>
<td></td>
</tr>
<tr>
<td>Visual impairment (18–48%)</td>
<td></td>
</tr>
<tr>
<td>Language disorders (37–64%)</td>
<td></td>
</tr>
<tr>
<td>Agitation, aggression, tantrums, attention deficits, memory deficits, inhibition or initiation deficits (23–59%)</td>
<td></td>
</tr>
</tbody>
</table>
practice, frequently proffer various speculative causation theories (described in prior sections) camouflaged as alternative or mimic diagnoses in child maltreatment cases. These medical witnesses run afoul of professional norms and standards and, when their arguments are repeated by the news media, create a grave public health risk by promulgating dangerous misinformation regarding safe infant and child care.

**What can be done to provide the court accurate information about the state of medical knowledge in AHT**

**The admissibility of expert evidence**

In current day jurisprudence, admissibility of medical or scientific expert testimony requires some judicial assessment of the “reliability” of that testimony. In some jurisdictions, the standard for assessing admissible expert testimony is the *Frye* standard (or whether a particular concept or methodology is “generally accepted” in the medical/scientific community); in others, it is a *Daubert* standard (where judges consider additional criteria other than just “general acceptance,” such as testability, peer review and publication and error rate). But in any legal jurisdiction, the medical precept that is considered “generally accepted” holds significant weight with courts. Unfortunately, courts are generally ill-equipped to measure the general consensus of physician thought on a particular concept, which makes them susceptible to more speculative theories unsupported by the medical evidence and medical literature. Thus, consensus statements present a unique opportunity to provide courts with a way to know general medical thought about a particular medical topic.

**Professional society consensus statements**

**Physician acceptance** Courts should assume that a consensus statement reflects general physician acceptance of a particular precept. Table 6 describes the rigorous process used to construct this type of statement. Thus, courts can be assured that practice promulgation of consensus statements has been vetted through a process that offers all members a way to contribute to the professional statements of that medical society.

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Process for developing a consensus statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Topic under society’s expertise needs clarification</td>
<td></td>
</tr>
<tr>
<td>b. Governing body of a society appoints individuals or a society’s committee with expertise on a subject to study the issue and write a statement</td>
<td></td>
</tr>
<tr>
<td>c. The appointed group (the writing group) may utilize experts from other medical subspecialties and other professional societies as consultants and authors</td>
<td></td>
</tr>
<tr>
<td>d. A draft document is created and reviewed by participating individuals,</td>
<td></td>
</tr>
<tr>
<td>e. The document, after modification by this input is sent to the governing body of the specific society for comments</td>
<td></td>
</tr>
<tr>
<td>f. With these comments, the writing group revises the document and submits to the governing body for approval</td>
<td></td>
</tr>
<tr>
<td>g. The governing body circulates the document to the society membership for comment and if necessary further revisions</td>
<td></td>
</tr>
<tr>
<td>h. After this comprehensive creation and review process is completed, the document is published</td>
<td></td>
</tr>
</tbody>
</table>

**Education of the courts** Professional consensus statements can influence the judicial process through interdisciplinary education. Courts need experts to provide general information about infant anatomy, imaging technologies and the interpretation of medical images and laboratory results. To perform their decision-making role, judges and juries must assess the weight of the medical literature and differentiate between persuasive evidence-based medical research and less persuasive or unpersuasive published work (e.g., opinion articles, single case studies or discredited articles). In AHT, pediatricians and pediatric subspecialist physicians can be crucial to a court’s accurate understanding of the relevant and reliable medical evidence.

Experts, through consensus statements, can also help courts identify the medical evidence that reflects scientific knowledge because it is supported by the evidence and has been generally accepted in the relevant field of pediatric medicine. By providing that medical information in a consensus statement, professional medical societies assist courts in identifying testimonial parameters for expert testimony and help judges and juries delineate evidence-based medical knowledge from fringe, speculative, or professionally irresponsible opinions.

**Accurate medical evaluation versus non-evidence-based opinions**

In cases involving an AHT diagnosis by one or more physicians, defense attorneys and their retained medical witnesses have increasingly challenged longstanding medical consensus that infant shaking can cause brain trauma. Typical defense arguments include: (1) a biased rush to judgment on the diagnosis of abuse; (2) exclusive diagnostic reliance on a “triad” of symptoms; (3) diagnosis by default; (4) an absence of neck injuries, proving AHT did not occur; (5) shifting scientific consensus; (6) an epidemic of copycat false convictions; and (7) the presumption that confession evidence consistent with infant injuries was coerced (in the two papers on confessions from France, in fact, the perpetrators were offered no reason to confess because leniency cannot be offered via French law) [31, 32]. These arguments are repeatedly raised in court despite the fact that they have never been empirically substantiated or are patently false.

There is a major flaw propagated in the few articles of those who deny SBS/AHT. It is the erroneous use of the terms “evidence-based medicine” and “systematic review” [200].
Because the suggestion that denialist views are supported by the evidence is likely to confuse judges and juries, we address two purported literature reviews: Donohoe in 2003 [201] and Lynoe et al. in 2017 [202]. Both articles are flawed by “(1) improper search and systemic review questions, (2) improper criteria for assessing bias and (3) inequitable application of quality of study assessment standards” [137, 203].

It is unprecedented that Donohoe’s “systematic review” chose to exclude the voluminous literature before 1999 despite the fact that AHT was well described by multiple authors worldwide and the incidence of the disease was quite similar worldwide [204]. In the final analysis, Donohoe used only 23 articles to reach his erroneous conclusions [201]. As Greeley [204] showed, evidence supporting the AHT medical diagnosis “clearly fits the Bradford Hill criteria for causation” [205]. Similarly, despite the vast medical literature, Lynoe et al. [202] chose to use only 30 publications. Narang et al. [203] revealed the severe prejudicial bias of the authors of the Lynoe et al. [202] study. Additional publications have also refuted the Lynoe report [206–210]. This alternative agenda has no role in true science and can result in infant harm through shaking and neglect, through avoidance of emergency medical intervention.

In contrast, a 2016 study published in The Journal of Pediatrics found a high degree of medical consensus that shaking a young child can cause subdural hematoma, severe retinal hemorrhage, coma or death [3]. The study, which surveyed 628 physicians at 10 leading U.S. children’s hospitals, found that 88% of physicians believe that SBS is a valid evidence-based diagnosis and 93% believe that the somewhat more comprehensive diagnosis of AHT is a valid evidence-based diagnosis [3].

**AHT is a medical diagnosis, not a legal finding of murder**

It is increasingly popular for defense lawyers to argue that AHT is a medical diagnosis of murder. This evocative courtroom hyperbole deliberately distorts the judicial process by mischaracterizing the physician expert’s role. The medical expert in a child abuse case plays just one role — to help the judge or jury answer the medical question of whether an infant’s injuries were most likely caused by abuse or they could be plausibly explained by a recognized disease or by one or more of the myriad hypothetical alternative causal explanations typically proffered by the defense. It is absurd to argue that a medical diagnosis proves murder. Medical expert testimony on the etiology of the injury cannot answer the two foundational legal questions of actus reus (Latin for guilty act) or mens rea (Latin for guilty mind). That is because, even after the factfinder decides that the medical evidence supports a finding that an infant’s injuries were inflicted, non-medical evidence is required to determine who committed the act and to determine the level of intent (e.g., knowing, reckless or negligent). “The debate surrounding AHT is neither scientific nor medical but legal” [204]. The denialists have tried to create a medical controversy where there is none.

The “diagnosis of murder” argument is obviously wrong because it falsely implies that medical opinion testimony, by its nature, resolves all legal issues. To cite an analogous example that disproves the argument’s premise, the toxicologist who testifies that the victim was poisoned does not diagnose murder because the court must still decide the actus reus (how was the poison ingested?) and the mens rea (was the victim’s poisoning accidental, negligent, reckless or intentional?).

Defense attorneys and few medical witnesses who promulgate scientifically unsubstantiated theories about abuse “mimics” in an effort to manufacture a scientific-sounding controversy run afoul of professional norms and standards, can distort the view of the relevant medical community, and create a grave public health risk by promulgating dangerous misinformation regarding safe infant and child care (i.e. infant shaking is safe). As professional medical societies continue to issue evidence-based consensus statements to help courts, the news media and the public to address these issues, we anticipate that they will also play a greater role in curbing and sanctioning members whose testimony impedes the goals of scientific, adjudicative and public health accuracy.

**Conclusions**

1. Abusive head trauma (AHT) is the current most appropriate and inclusive diagnostic term for infants and young children who suffer from inflicted intracranial and associated spinal injury. This does not negate the mechanisms of shaking or shaking with impact as a significant mechanism of injury but merely indicates that the term “shaken baby” is not all-inclusive.
2. Lack of history, changing history or the incompatibility of history (i.e. short falls) with the severity of injury raise concerns for possible AHT.
3. Relatively few infants with AHT have isolated intracranial injury without retinal hemorrhages, fractures or other manifestations of child abuse. These children need a comprehensive evaluation to rule out other diseases. However, isolated intracranial injuries occur in a small percentage of children with AHT.
4. No single injury is diagnostic of AHT. A compilation of injuries most often including SDH, complex retinal hemorrhage and/or retinoschisis, rib, metaphyseal or other fractures and soft-tissue injury leads to the diagnosis.
5. Each infant suspected of suffering AHT must be further evaluated for other diseases that might present with similar findings. The question to be answered is, “Is there a medical cause to explain the findings or did this child suffer from inflicted injury?”
6. There is no reliable medical evidence that the following processes cause the constellation of injuries associated with AHT: cerebral sinovenous thrombosis, isolated hypoxic–ischemic injury, lumbar puncture and dysphagic choking/vomiting. There is no reliable evidence to support speculation that long-term consequences of birth-related subdural hematoma can result in later collapse, coma or death from acute rebleeding into a previously asymptomatic chronic subdural hematoma. In addition, subdural hematoma is uncommon in the setting of benign enlargement of the subarachnoid space, and when present, AHT should be considered in the differential diagnosis.

7. After medical diagnosis, in many hospitals a multidisciplinary team provides comprehensive assessment and services to the family, based on consideration of all the facts.

8. There is no controversy about the methodology used to diagnose AHT as a medical disease.

9. AHT is a medical diagnosis unrelated to the legal determination by a judge or jury of a charge of murder. The term “triat” is a legal convention that falsely mischaracterizes a complex AHT diagnosis process.

10. A professional medical society’s consensus statement educates judicial factfinders, the news media and the public about “general acceptance,” what is accurate medical information and what is non-evidence, speculative or professionally irresponsible etiological hypotheses.

11. The professional societies’ consensus statement on AHT should help the court recognize unsubstantiated medical expert testimony.

Acknowledgments We are grateful to the fellows at Children’s Hospital of Philadelphia who reviewed the references for this manuscript (Andrew J. Degnan, Rachelle Durand, Edward Fenlon, Ami Gokli, Aditi Hendi, James Hogan, Fang Lu, Ian Mills, Christy Pomeranz, Jordan Rapp and Michele Retrouvey).

Compliance with ethical standards

Conflicts of interest Dr. Narang has been paid as an expert consultant/witness in cases of abusive head trauma. Drs. Choudhary, Servaes, Christian, Hedlund, Dias, Nelson, Palasis, Rossi and Offiah provide medical–legal expert work in child abuse cases.

References

Pediatr Radiol


