Akron Law Review

Volume 54 | Issue 1 Article 1

Emergent Neurotechnologies and Challenges to Responsibility Frameworks

Laura Cabrera

Jennifer Carter-Johnson

EMERGENT NEUROTECHNOLOGIES AND CHALLENGES TO RESPONSIBILITY FRAMEWORKS

Laura Cabrera and Jennifer Carter-Johnson**[1]*

* Laura Cabrera is an Associate Professor of Neuroethics at the Center for Neural Engineering, Department of Engineering Science and Mechanics at The Pennsylvania State University.

** Jennifer Carter-Johnson is the Associate Dean for Academic Affairs, Associate Professor of Law, Director of the Intellectual Property and Entrepreneurship Clinic, and the Co-Associate Director of the Intellectual Property, Information and Communication Law Center at Michigan State University College of Law.

1. This article has benefitted from comments from the participants at the Wiet Life Science Law Scholars Workshop hosted by Loyola University Chicago. All errors are the authors'.

I. INTRODUCTION

"It is not my fault; it is my brain implant which made me do it." Some scholars have argued that this could become a common strategy: defendants might argue that as the result of a defective brain implant, an autonomous brain implant, or someone hacking into their implant, they should not be held responsible, or at least not *fully* responsible.

In the past few years, a neuroscientific revolution has been underway. Neuroscience has rapidly increased our knowledge of the functioning of the human brain, providing us with an insight into the mental processes underpinning human behavior. This explosion of interest in neuroscience has resulted in the development of many neurofields: from *neuro*aesthetics to *neuro*economics and *neuro*marketing. But as we learn more about the brain, we also learn more about human thought and motivations. These new understandings and knowledge about the functioning of the human brain are of great relevance to ethics and law, given that these are disciplines primarily concerned with the normative dimension of human behavior. That is why ethicists and legal scholars have been interested in the impact of neuroscientific advances, resulting in the rapid development of neuroethics[2] and neurolaw.[3]

Neuroethics deals with the ethical issues in the design and conduct of neuroscientific studies and includes topics already known in bioethics, such as informed consent, privacy, and risk assessment, but there are other

2. *See generally* NEIL LEVY, NEURO ETHICS (2007); Martha J. Farah, *Neuroethics: The Ethical, Legal, and Societal Impact of Neuroscience*, 63 ANN. REV. PSYCHOL. 571 (2012); NEUROETHICS (Martha J. Farah ed., 2010); Judy Illes & Stephanie J Bird, *Neuroethics: A Modern Context for Ethics in Neuroscience*, 29 TRENDS NEUROSCIENCE 511 (2006).

3. *See generally* Nicole A. Vincent, *Neurolaw and Direct Brain Interventions*, 8 CRIM. L. & PHIL. 43 (2012); Gerben Meynen, *Neurolaw: Neuroscience, Ethics, and Law. Review Essay*, 17 ETHICAL THEORY MORAL PRAC. 819 (2014); F. X. Shen, *Law and Neuroscience 2.0*, 48 ARIZ. ST. L.J. 1043 (2016).

topics that are more specific to the field of neuroscience.[4] In this regard, neuroethics is truly novel as it attempts to investigate the impact that our growing understanding of brain function, and novel technologies to manipulate brain function, may have on our ethical, social, and philosophical conceptions. Neuroethics includes issues like personal identity, freedom and responsibility, consciousness, and the mind body problem.[5] Neurolaw attempts to explore the influence that neuroscience's discoveries may have on legal rules and court decisions, in particular the use of evidence resulting from neurotechnologies.[6]

As neurotechnology advances and opens novel opportunities for monitoring and controlling brain function, there is uncertainty on how the law should cope with such advancements. 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 and criminal law, and legal philosophy.

This article explores a particular type of neuro-intervention, namely brain implants, with a focus on deep brain stimulation (DBS), and the possible challenges these might bring to current legal and ethical frameworks of responsibility. Clinically available brain implants have advanced functionalities, with some able to adjust stimulating parameters while others are constantly monitoring brain signals and adapting their stimulation parameters based on that data.[7] Some brain implants, such as those for DBS,[8] are well-accepted treatments for movement disorders,[9] and their use as treatment options for various psychiatric disorders is being explored.[10]

In this article, we delve into these challenging questions, providing an in-depth overview of the different issues that are raised. In Part II, we

4. Adina Roskies, *Neuroethics for the New Millenium*, 35 NEURON 21 (2002).

5. *See generally* sources cited *supra* note 1.

6. *See* sources cited *supra* note 2.

7. Meng-Chen Lo and Alik S. Widge, *Closed-Loop Neuromodulation Systems: Next generation Treatment for Psychiatric Illnesses*, 29 INT'L REV. PSYCHIATRY 191 (2017).

8. In what follows, we will be using the terms brain implants and DBS interchangeably, though is important to be aware that DBS is a type of brain implant.

9. Paul S. Larson, *Deep Brain Stimulation for Movement Disorders*, 11 NEUROTHERAPEUTICS 465 (2014); Jeff M. Bronstein, Michele Tagliati, Ron L. Alterman, Andres M. Lozano, Jens Volkmann, Alessandro Stefani, Fay B. Horak, Michael S. Henderson, Marwan I. Hariz, Roy A. Bakay, Ali Rezai, William J. Marks, Jr., Elena Moro, Jerrold L. Vitek, Frances M. Weaver, Robert E. Gross & Mahlon R. DeLong, *Deep Brain Stimulation for Parkinson Disease*, 68 ARCHIVES NEUROLOGY 165 (2011).

10. *See* Ilse Graat, Martijn Figee & Damiaan Denys, *The Application of Deep Brain Stimulation in the Treatment of Psychiatric Disorders*, INT'L REV. PSYCHIATRY 178 (2017); Helen Shen, *Tuning the Brain*, 507 NATURE 290 (2014).

will introduce a number of scenarios with added variations involving a brain implant and undesired behavior resulting in another person being harmed. Part IV discusses the general frames of moral responsibility and Part V touches on the legal considerations.

II. DBS CHALLENGES RESPONSIBILITY FRAMEWORKS

A. *DBS Changes*

DBS involves the surgical implantation of at least one electrode in the brain, and the implantation of a pulse generator (sometimes called a "brain pacemaker") under the patient's clavicle or in the abdomen, which controls the settings of the brain implant (e.g. voltage and frequency).[11] The pulse generator is carefully programmed for each patient to deliver electrical impulses to specific targets in the brain. DBS is considered more precise than earlier forms of psychosurgery, due to neuroimaging and other tools used to help guide the implantation of the electrodes within a millimeter of their target. Unlike lesioning methods, where a part of the brain is destroyed or removed, DBS is a neuromodulation approach, in which electrical pulses modulate brain activity in targeted areas. It is because of this feature that DBS is considered both adjustable and reversible,[12] as the electrical stimulation can easily be adjusted or turned off or on. However, given that the procedure still requires the implantation of electrodes deep in the brain, the possibility of unintended lesioning that is not reversible should not be fully discarded.[13]

While DBS aims to treat the underlying pathophysiology of neurological and psychiatric disorders, there are scenarios where brain implants may influence personality and affect an individual's behavior in undesired ways.[14] There are already real scenarios where brain implants have influenced an individual's perception of the world and behavior in

11. Joel Perlmutter & Jonathan Mink, *Deep Brain Stimulation*, 29 ANN. REV. NEUROSCIENCE 229, 230–31 (2006).

12. Sabine Müller, Rita Riedmuller & Ansel van Oosterhout, *Rivaling Paradigms in Psychiatric Neurosurgery: Adjustability Versus Quick Fix Versus Minimal-Invasiveness*, 9 FRONTIERS INTEGRATIVE NEUROSCIENCE 1, 2 (2015).

13. *See* Jennifer Mundale, *Reversibility and Deep Brain Stimulation*, 3 J. COGNITION & NEUROETHICS 97 (2016). For further discussion on the reversibility of DBS see Jonathan Pugh, *No Going Back? Reversibility and Why It Matters for Deep Brain Stimulation*, 45 J. MED. ETHICS 225 (2019).

14. Robyn Bluhm, Laura Y. Cabrera & Rachel McKenzie, *What We (Should) Talk About When We Talk About Deep Brain Stimulation and Personal Identity*, 13 NEUROETHICS 289 (2020).

unexpected ways.[15] There is, for example, the scenario of Mr. B, a man who received DBS as a treatment for his severe obsessive-compulsive disorder. Mr. B had never been a music lover until, under DBS, he "developed a distinct and entirely novel music preference for Johnny Cash . . . ."[16] When the device was turned off, Mr. B no longer liked Johnny Cash.[17] Another scenario involved an epilepsy patient who received DBS as part of an investigation to locate the origin of his seizures, and who experienced hallucinations during the stimulation.[18]

While these two scenarios did not result in harm to the patient or others, we can think of scenarios when actions of people with brain implants might result in harm to the patient or others. This raises a number of ethical and legal questions. For example, if brain-implant-induced changes in personality result in undesirable or deviant behaviors that cause harm, who (or what) is responsible? Is the person with the implant responsible? Can the implant be held responsible? Or can the engineer that developed, or the company that manufactured, the implant be held responsible? Ethical and legal questions related to responsibility are likely to become more salient as implants increasingly include capabilities for wireless communication and remote monitoring systems, which introduce potential cybersecurity concerns involving malicious interference with battery life or essential programming functions. What happens if someone illicitly accesses and manipulates someone else's brain implant ("malicious brain hacking")?[19] Who is responsible for the actions of the person whose implant is been hacked?

15. Nsikan Akpan, *Deep Brain Stimulation Triggers Hallucinations*, SCIENCE (Apr. 16, 2014, 1:45 PM), https://www.sciencemag.org/news/2014/04/scienceshot-deep-brain-stimulation-triggers-hallucinations [https://perma.cc/9376-XQDN]. *See* A.F. Leentjens, V. Visser-Vandewalle, Y. Temel & F.R. Verhey, *Manipulation of Mental Competence: An Ethical Problem in Case of Electrical Stimulation of the Subthalamic Nucleus for Severe Parkinson's Disease*, 148 NEDERLANDS TIJDSCHRIFT VOOR GENEESKUNDE, 1394 (2004); Frederik Gilbert, *Deep Brain Stimulation For Treatment-Resistant Depression: Postoperative Feeling of Self-Estrangement, Suicide Attempt, and Impulsive-Aggressive Behaviors*, 6 NEUROETHICS, 473 (2013).

16. Mariska Mantione & Damiaan Denys, *A Case of Musical Preference for Johnny Cash Following Deep Brain Stimulation of the Nucleus Accumbens*, 8 FRONTIERS BEHAV. NEUROSCIENCE 1 (2014).

17. Nicky Woolf, *Man Develops Powerful Love of Johnny Cash Following Deep Brain Stimulation*, GUARDIAN, May 27, 2014, https://www.theguardian.com/music/2014/may/27/johnny-cash-deep-brain-stimulation-urge-listen [https://perma.cc/6DEJ-V54F].

18. Akpan, *supra* note 15.

19. Marcello Ienca & Pim Haselager, *Hacking the Brain: Brain-Computer Interfacing Technology and the Ethics of Neurosecurity*, 18 ETHICS & INFO. TECH. 117, 117 (2016). Ienca and Haselager use the term "brain hacking" to refer to activities that directly influence neural computation in the users of neurodevices in a manner that resembles how computers are hacked in computer crime. *See also* Laurie Pycroft, Sandra G. Boccard, Sarah L.F. Owen, John F. Stein, James J. Fitzgerald, Alexander L. Green & Tipu Z. Aziz, *Brainjacking: Implant Security Issues in Invasive*

B. DBS Situational Possibilities

1. Unexpected personality change, Scenario 1

A 62-year-old Dutch man, Mr. D, begins experiencing manic episodes approximately three years after being implanted with an open-loop DBS system for treatment of Parkinson's disease. Therapy with psychiatric medication fails to control the symptoms, which include megalomania and impulsivity. His psychological condition eventually degrades to the point where he is no longer competent to care for himself, and he is admitted to a psychiatric hospital. Adjustment of his DBS system causes the manic symptoms to abate, but in the absence of DBS Mr. D's Parkinsonism is so severe that he becomes bedridden. Ultimately, Mr. D has to choose between a nursing home, where he would be bedridden but coherent, and a psychiatric ward, where he would be mobile, but manic. He chooses the latter.[20]

In this scenario, there is harm to the patient due to side effects of the DBS treatment, but no harm to anyone else. Under what circumstances should the DBS manufacturer be liable for the harm caused to Mr. D? The cost of psychiatric treatment is not minimal and may not be covered by insurance in the same way as the Parkinson's disease symptoms.

2. Unexpected personality change, Scenario 2

Returning to Mr. B, we might imagine a scenario where he has further issues arising, albeit indirectly, from his DBS treatment. Mr. B loves Johnny Cash, except when he doesn't. Mr. B received DBS as a treatment for his severe obsessive-compulsive disorder. He had never been a music lover until, under DBS, he developed a liking for Johnny Cash. When the device is turned off, the preference disappears.[21] Let us think about the potential consequences: while using the device, Mr. B spends thousands of dollars collecting Johnny Cash's music and memorabilia—ending up hopelessly in debt. When Mr. B loses his job and home, foreclosure threatens due to his lack of savings, and he decides to rob a bank rather than sell his Johnny Cash collection.

Neuromodulation, 92 WORLD NEUROSURGERY 454 (2016); Mark N Gasson & Bert-Jaap Koops, *Attacking Human Implants: A New Generation of Cybercrime*, 5 L., INNOVATION & TECH. 248 (2013).

20. Felicitas Kraemer, *Authenticity or Autonomy? When Deep Brain Stimulation Causes a Dilemma*, 39 J. MED. ETHICS 757, 757–58 (2013) (adapted from Leentjens et al., *supra* note 15).

21. Woolf, *supra* note 17.

Who is liable for the harm due to the bank robbery? The DBS treatment did not directly force Mr. B to rob the bank, but it set up the love for Johnny Cash that led to depleting savings accounts and thus the robbery. Would Mr. B have robbed the bank absent the DBS treatment? Did DBS change his state of mind or just his circumstances? What if, without DBS, Mr. B's savings would not have been enough to save his home due to other costs associated with his disease?

3. Malfunction Scenario

Imagine that Ms. Q is driving one day and has a sudden urge to swerve into a bus stop where several people are standing. As a result, she ends up damaging the bus stop and injuring several people. During the investigation, police find that Ms. Q has a brain implant to treat her Parkinson's disease. This implant malfunctioned at the time the urge occurred. Who is liable for the damage to the bus stop?

Common causes for DBS malfunction might include electrode-related failure (electrode migration, electrode fracture), unit malfunction (battery failure or component malfunction), and malfunction related to exposure to high voltage electricity or high-intensity microwaves. These malfunctions can result in rebound symptoms,[22] as well as worsening of symptoms, suicidality, mood disturbances, and panic attacks.[23]

4. Closed-loop AI system

Think again of the hypothetical scenario of Ms. Q. This time assume that during the police investigation no malfunction of the device is found. However, it is revealed that Ms. Q had a closed-loop brain implant, as part of her DBS treatment, rather than an open-loop implant.

Closed-loop brain implants have been in use for certain types of treatment-refractory epilepsy (Neuro Pace RNS System). In addition, closed-loop DBS systems just gained Food and Drug Administration (FDA) approval to be used for patients with Parkinson's disease, essential

22. Marwan I. Hariz & F. Johansson, *Hardware Failure in Parkinsonian Patients with Chronic Subthalamic Nucleus Stimulation is a Medical Emergency*, 16 MOV. DISORD. 164 (2001); François Alesch, *Sudden Failure of Dual Channel Pulse Generators*, 20 MOV. DISORD. 64 (2004); Richard G Bittar, John Yianni, ShouYan Wang, Xuguang Liu, Dipankar Nandi, Carole Joint, Richard Scott, Peter G. Bain, Ralph Gregory, John Stein & Tipu Z. Aziz, *Deep Brain Stimulation for Generalised Dystonia and Spasmodic Torticollis*, 12 J. CLINICAL NEUROSCI. 12, 14 (2005); A.K. Vora, H. Ward, K.D. Foote, W.K. Goodman & M.S. Okun, *Rebound Symptoms Following Battery Depletion in the NIH OCD DBS Cohort: Clinical and Reimbursement Issues*, 5 BRAIN STIMULATION 599 (2012).

23. Vora et al., *supra* note 22.

tremor, dystonia, epilepsy, or OCD.[24] The direct way in which these implants are linked and interact with the human brain can make the source of an act difficult to identify.[25]

Human accountability for harms caused by the use of closed-loop devices might occur in several ways. For example, if a form of "veto" is built into the system,[26] then a person could be held accountable for failing to exercise the veto. But what happens when we have fully autonomous implants using machine learning algorithms? An interesting parallel here is that of autonomous cars or autonomous weapon systems with evolving algorithms controlling their actions. Here, the issue is that the implant's decisions are not preprogrammed, but rather are based on acquired experience.[27] In these scenarios, the greater degree of autonomy—being able to respond and adapt to the environment in unpredictable and intelligent ways—opens up a potential "responsibility gap."[28]

5. Brain hacking

To complicate matters more, what happens if someone illicitly accesses and manipulates someone else's brain implant ("malicious brain hacking")? A brain implant remotely compromised by a third party to commit an offense such as an assault would make it even more challenging to determine causality and obtain proof in legal proceedings. Who is responsible for the actions of the person whose implant has been hacked? The possibility of "brain hacking" is a serious consideration. At present, examples of ways in which brain hacking can affect user's behaviors and thoughts is by disrupting the functionality of the brain implant (e.g. turning off the stimulation, changing the stimulation

---

24. *FDA Approves First-Of-Its-Kind Percept™ PC Neurostimulator with BrainSense™ Technology*, MEDTRONIC NEWSROOM (June 25, 2020), https://newsroom.medtronic.com/news-releases/news-release-details/fda-approves-first-its-kind-percepttm-pc-neurostimulator [https://perma.cc/7CVZ-UJN8].

25. *See* Eliza Goddard, *Deep Brain Stimulation Through the "Lens of Agency": Clarifying Threats to Personal Identity from Neurological Intervention*, 10 NEUROETHICS 325 (2017). Goddard discusses different ways in which the threat to identity has been framed, arguing that "the ethically salient issue from DBS is impacts on autonomous agency—whether one's actions and beliefs are one's own". *Id.* at 326.

26. *See* Jens Clausen, Eberhard Fetz, John Donoghue, Junichi Ushiba, Ulrike Sporhase, Jennifer Chandler, Niels Birbaumer & Surjo R. Soekadar, *Help, Hope, and Hype: Ethical Dimensions of Neuroprosthetics*, 356 SCIENCE 1338, 1338 (2017).

27. Andreas Matthias, *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, 6 ETHICS & INFO. TECH. 175, 176 (2004) (quoted in Mark A Chinen, *The Co-Evolution of Autonomous Machines and Legal Responsibility*, 20 J.L. TECH. 1 (2016)).

28. Matthias, *supra* note 27.

parameters).[29] In the future, brain hacking could also involve illicit access to brain information that can be used in a manner that resembles the threats caused by computer hacking, such as blackmailing individuals, selling the information to third parties, or generating undesired behaviors.[30] These malicious modifications may threaten the person with the implant by causing a DBS device to operate, for example, out of safe parameters. Moreover, the forced intrusion into and alteration of a person's neural processes poses an unprecedented threat to that person's mental integrity because these activities can "limit and constrain their behavior, generate emotional responses such as panic, fear, and psychological distress, and leave traumatic memories."[31] While these neurosecurity threats are still mostly theoretical, they are increasingly likely to emerge as software and remote monitoring become embedded in more medical devices.[32]

All the above scenarios invite us to think about moral and legal responsibility as technology advances. Is Ms. Q solely responsible for her actions? Can we attribute any blame to the DBS device, in particular in the scenario of closed-loop systems? What about the engineers who designed it or the company that manufactured it? The neurosurgeon who implanted it or the neurologist who programmed the device parameters?

Lawyers, philosophers, and ethicists have labored in a parallel fashion to define the conditions under which individuals are to be judged legally and morally responsible for their actions.[33] The scenario of brain implants and questions about responsibility are of particular interest because the brain is generally regarded as the center of control, rational thinking, and emotion. It is the organ orchestrating people's actions and behaviors. As such, the brain is key to agency, autonomy, and responsibility.

---

29. Ienca & Haselager, *supra* note 19, at 119. *See also* Pycroft et al., *supra* note 19.

30. Ienca & Haselager, *supra* note 19, at 120; Pycroft et al., *supra* note 19, at 456. Pycroft and colleagues summarize under Table 1 of their article different forms of attack and the potential harms they can cause to individuals.

31. *See generally* Marcello Ienca & Roberto Andorno, *Towards New Human Rights in the Age of Neuroscience and Neurotechnology*, 13 Life Sci. Soc. Pol'y 1 (2017); Ienca & Haselager, *supra* note 19, at 119.

32. Pycroft et al., *supra note* 19, at 454 (Neurosecurity refers to defense mechanisms protecting brain devices); *See* Daniel B. Kramer & Kevin Fu, *Cybersecurity Concerns and Medical Devices*, JAMA 1–2 (2017); Gasson & Koops, *supra* note 19; Pycroft et al., *supra* note 19.

33. *See, e.g.*, ARISTOTLE, NICOMACHEAN ETHICS 1 (W.D. Ross trans., Internet Classics Archive) (c. 384 B.C.E.) Part II Section 4; T. Wardlaw Taylor, Jr., *The Law and Responsibility*, 7 PHIL. REV. 276 (1898); Andrew Eshleman, *Moral Responsibility*, *in* STAN. ENCYCLOPEDIA PHIL. (Edward N. Zalta ed., 2014); John Martin Fischer, *Recent Work on Moral Responsibility*, 110 ETHICS 93 (1999).

In what follows, we will explore the ethical and legal challenges of responsibility related to DBS and brain implants in general.

III. ETHICAL FRAMEWORK

Historically, moral and legal responsibility have predominantly focused on the autonomous individual, i.e., someone with the capacity to deliberate or act on the basis of one's own desires and plans, free of distorting external forces.[34] However, with modern technological advances, many "hands" may be involved in the operation of these brain implants, including artificial intelligence agents directly influencing the brain. This external influence raises questions about the degree to which someone with an implant can control actions and behaviors. If brain implants influence someone's decisions and behaviors, do they undermine the person's autonomy? If autonomy is undermined, can we attribute responsibility to the individual? What makes a certain agent responsible for a certain event? Can we ever ascribe responsibility to an artificial intelligence-based brain implant? Responsibility is a complex concept used both in law and in moral philosophy to describe the attribute or state of being responsible. It is generally associated with "getting the credit or blame for acts or decisions," being "liable to be called to account as the primary cause, motive, agent," or being "liable to legal review or in case of fault to penalties."[35] These definitions reflect various understandings of moral and legal responsibility common in the West, as well as the ways in which the concept is intertwined with, and overlaps with, notions like accountability, liability, blameworthiness, and causality.

A. *Moral responsibility*

There are ongoing debates on what sets a particular form of responsibility, such as moral responsibility, apart from other kinds of responsibility. A generally accepted view on moral responsibility says that a person is *morally* responsible when her voluntary actions lead to *morally* significant outcomes that would make her suitable to praise or blame. The causal connection between the person and the outcomes of voluntary actions is one of the main conditions to ascribe moral responsibility

34. Lewis Hinchman, *Autonomy, Individuality, and Self-Determination, in* WHAT IS ENLIGHTENMENT? EIGHTEENTH-CENTURY ANSWERS AND TWENTIETH-CENTURY QUESTIONS 488 (James Schmidt ed., 1996).

35. *Responsible*, MERRIAM WEBSTER DICTIONARY, http://www.merriam-webster.com/dictionary/responsible [https://perma.cc/YL9Q-YW6Y].

(volitional condition).[36] Another important condition is that the agent must have knowledge of what the agent is doing in order to be able to consider the possible consequences of a given action (epistemic condition). Thus, if someone could not have known that his or her actions would lead to a harmful event, we tend to mitigate responsibility (and in some cases even excuse the person). This condition of responsibility goes back to Aristotle, who argued that since ignorance is a form of involuntary action, it could remove or diminish a person's moral responsibility.[37] Another important criterion is the ability to freely choose to act in a certain way, so one cannot be held responsible for actions determined by outside forces.[38]

1. Causal contribution

In the scenarios we are exploring here, it is not clear the impact on decision making and behavior such brain implants can have on an individual. Holding a person responsible for what someone else has done, or an event she has no control over, is generally seen as unfair and unjustified.[39] Brain implants complicate the causal connections between a person's actions and the consequences. Can we hold the person morally responsible for an action that was caused by the brain implant? Technologically-mediated, undesired events are usually the result of an accumulation of errors, biases, or failures to notice or do something of various individuals (and in some cases even algorithms) involved in the development, function, use, maintenance, and regulatory oversight of a brain implant. Take the hypothetical scenario with Ms. Q. The crash with the bus stop could have resulted from the combination of a number of factors, including battery depletion, hardware errors, programming errors, inadequate testing of programming settings, overconfidence in programming settings or machine learning algorithms, and inadequate investigation of adverse events or accident reports. In addition, there are

36. *See* Harry G. Frankfurt, *Necessity and Desire*, 45 PHIL. & PHENOMENOLOGICAL RES. 1, 3 (1984); Harry G. Frankfurt, *Freedom of the Will and the Concept of a Person*, 68 J. PHIL. 5, 10 (1982).

37. ARISTOTLE, *supra* note 33, at Book III Section 1.

38. *Id.* Aristotle pointed to all of these conditions: "[I]f the acts that are in accordance with the virtues have themselves a certain character it does not follow that they are done justly or temperately. The agent also must be in a certain condition when he does them; in the first place he must have knowledge, secondly he must choose the acts, and choose them for their own sakes." *See* Adina Roskies, *Neuroscientific Challenges to Free Will and Responsibility*, 10 TRENDS COGNITIVE SCI. 419 (2006). Contemporary views about freedom of choice have been contested in light of new neuroscientific evidence, however as Roskies argues "they are not likely to affect practical judgments of moral responsibility." *Id.* at 422.

39. Fischer, *supra* note 33, at 95.

multiple individuals whose actions (or inactions) could have shaped the outcome, including engineers, device manufacturers, clinicians, users, and even policymakers.

The fact that a given outcome involves multiple agents (human and non-human), each one contributing in different ways to the outcome, gives rise to what is known as the problem of "many hands."[40] The more agents (human and non-human) contribute to a given outcome, the more difficult is to determine who or what was responsible.[41] All of this complicates tracing the series of events that led to incidents like those presented in the above scenarios. Was the algorithm underlying the device functioning without errors? Were the manufactured parts of the device working within specifications for medical devices? Did the clinician adequately implant and program the device in the brain target? Did patients using the device follow the indications on things to do or not to do with the device implanted?

Another problem is the temporal and physical distance. The engineers of an automated brain implant, for example, make choices ahead of time determining how implants will act, but they rarely will see how these decisions impact patients with those implants. As Coeckelbergh argues, we are in a situation of "epistemic opacity" as "between our actions and the consequences of our actions lies a complex world of relationships, people, things, time and space."[42] Moreover, technologies are not isolated instruments; they are "socially constructed and society shaping."[43] In this regard, brain implants are one more example of complex technological systems complicating the task of attributing responsibility, both from a backward-looking and forward-looking perspective.[44]

---

40. *See generally* Ibo van de Poel, Lamber Royakkers, and Sjoerd D. Zwart, MORAL RESPONSIBILITY AND THE PROBLEM OF MANY HANDS (2015). Van de Poel and colleagues define the problem of many hands as "undesirable outcomes in collective setting for which it is hard or even impossible to hold an individual or organization . . . responsible" *Id.* at 4. Floridi also describes the problem of the "invisible hand" which results from the "systemic interactions among multiagent systems (comprising several agents, not all necessary human)." Luciano Floridi, *Distributed Morality in an Information Society*, 19 SCI, & ENGINEERING ETHICS 727, 728 (2012).

41. Van de Poel et al., *supra* note 40; Katinka Waelbers, *Technological Delegation: Responsibility for the Unintended*, 15 SCI. & ENGINEERING ETHICS 51, 52–53 (2009).

42. Waelbers, *supra* note 41, at 52; MARK COECKELBERGH, HUMAN BEING @ RISK, 103 (2013).

43. Thomas P. Hughes, *The Evolution of Large Technological Systems*, *in* THE SOCIAL CONSTRUCTION OF TECHNOLOGICAL SYSTEMS 45 (Wiebe E. Bijker, Thomas P. Hughes & Trevor Pinch eds., 2012).

44. Van de Poel et al., *supra* note 40, at 5.

2. Considering the consequences

Another important aspect of moral responsibility is the capacity of an agent to consider and deliberate about the consequences of her action. It is not well-established how much brain implants shape how people perceive and experience the world, or whether this affects in any relevant way the described condition for responsibility.

Brain implants are complex technologies, with users and clinicians often having only a partial understanding of the assumptions, models, and theories on which these implants operate. These issues become more salient when thinking of closed-loop implants with machine learning capabilities, where even the developers might not fully understand what lies behind the decision-making tree, or when thinking of capabilities that increase the likelihood of hacking a brain implant.[45] As philosopher John Ladd nicely captures it, "[t]echnology has created new modes of conduct and new social institutions, new vices and new virtues, new ways of helping and new ways of abusing other people".[46]

3. Freedom to act

Another important condition for the attribution of moral responsibility is the freedom to act, which in moral philosophy is often discussed as a person having free will or autonomy.[47] We hold individuals morally responsible when they have the capacity to control their behavior, on the basis of their own authentic reasons and motivations.[48] That is why we generally do not hold people responsible if they are manipulated or forced to take a particular action. Brain implants might affect the decisions that someone makes, and how he or she makes them, by facilitating and enabling particular human cognitive processes, actions, or attitudes while constraining and inhibiting others. As Verbeek argues, "technological artifacts are not neutral intermediaries but actively coshape people's being

45. *See* Ienca & Haselager, *supra* note 19; Pycroft et al., *supra* note 19; Eduard Marin, Dave Singelée, Bohan Yang, Vladimir Volski, Guy A.E. Vandenbosch, Bart Nuttin & Part Preneel, *Securing Wireless Neurostimulators*, PROC. EIGHTH ACM CONF. ON DATA & APPLICATION SECURITY & PRIVACY 287 (2018).

46. J. Ladd, *Computers and Moral Responsibility: A Framework for an Ethical Analysis, in* THE INFORMATION WEB: ETHICAL AND SOCIAL IMPLICATIONS OF COMPUTER NETWORKING 207, 210–11 (C. C. Gould ed., 1989).

47. Fischer, *supra* note 33.

48. *Id.* Fischer argues that control in moral responsibility is more about the possibility of an agent "to select from among various genuinely open paths the world might take," and not necessarily as control in the sense of self-governance that theories of autonomy often imply. *Id.* at 99.

in the world: their perception and actions, experience, and existence."[49] For instance, in the scenario of patient Q undergoing DBS, acts conducted as a result of hallucinations challenge this condition of moral responsibility.

There is also a lack of consensus on the conditions that enable individuals to act freely. For some it is rationality, for others it is intentionality or emotion, for others it is a combination of all of these. Recent neuroscience research has questioned whether human beings really act out of free will.[50] While we are not going to delve into a discussion of free will here, we do want to discuss a connected concept, autonomy.

The concept of autonomy within bioethics and philosophy is not without ambiguity,[51] but it is generally taken to be the capacity of someone to deliberate or act on the basis of one's own desires and plans and not as the product of manipulative or distorting external forces.[52] Research has shown how easy is to manipulate, control, or influence individuals, both by external forces (such as peer pressure) or internal ones (such as addictions or mental problems). However, direct brain interventions seem unique in that they surpass the conscious awareness of individuals to even decide on whether to follow a set of actions or not.[53] In this regard, brain implants add an additional layer of complexity in determining whether an individual has acted freely or autonomously, or if the implant had inserted enough influence to raise a question about the freedom of individual choice and behavior.

If the device is not controlled directly by the agent, like in the scenario of someone brain hacking the device, that seems like a classic case of manipulation, where we have anomalous and unusual causation.

49. Peter-Paul Verbeek, *Materializing Morality: Design Ethics and Technological Mediation*, 31 SCI., TECH. & HUM. VALUES 361, 364 (2006).

50. Roskies, *supra* note 38, at 419; *See also* Giuseppe Sartori, Silvia Pellegrini & Andrea Mechelli, *Forensic Neurosciences: From Basic Research to Applications*, 24 CURRENT OPINION NEUROLOGY 371, 371 (2011); Kerri Smith, *Taking Aim at Free Will*, NATURE 1 (2011); Patrick Haggard, *Human Volition: Towards a Neuroscience of Will*, 9 NATURE REV. NEUROSCI. 934 (2008).

51. *See, e.g.*, W. Glannon, *Neuromodulation, Agency and Autonomy*, 27 BRAIN TOPOGRAPHY 46 (2012); Frédéric Gilbert, *A Threat to Autonomy? The Intrusion of Predictive Brain Implants*, 6 AJOB NEUROSCI. 4 (2015); B. Jennings, *Reconceptualizing Autonomy: A Relational Turn in Bioethics*, HASTINGS CTR. REPORT (2016).

52. *See* Hinchman, *supra* note 34; S. Buss, *Personal Autonomy*, THE STAN. ENCYC. OF PHIL. (2002).

53. Vincent, *supra* note 3; Jan C. Bublitz and Reinhard Merkel, *Crimes Against Minds: On Mental Manipulations, Harms and Human Right to Mental Self-Determination*, 8 CRIM LAW PHILOS. 51 (2014). Bublitz and Merkel argue "direct interventions change the cognitive machinery itself." *Id.* at 69.

In the scenario of brain hacking, the manipulator bypasses the agency-relevant capacities of the agent, substantially undermining individual autonomy. This is important because in Western jurisprudence the capacity for voluntary control over one's actions is considered a requisite for individual legal liability regarding those actions.[54] Furthermore, as implants become more common, it is plausible that we are going to confront many more cases of anomalous causation, blurring the lines between ordinary agency and cases of manipulation.The difficulty of ascribing responsibility in light of the ways in which brain implants, such as the ones discussed here, muddle the conditions for it, is an indicator of the limitations of conventional ethical frameworks in dealing with the question of moral responsibility. While the nature of technology is relevant to the responsibility arrangements, such arrangements are, in the end, "socially constituted through the norms and expectations of particular activities and contexts."[55] In the case of brain implants societal attributions of responsibility are also impacted by the intentions and actions of other human agents involved in the design, manufacture, and programming of the device. Thus, acting with brain implants may require a different kind of analysis as to who can be held responsible and what it means to be morally responsible.

B. Shared responsibility

Moral responsibility has predominantly been about human action and its intentions and consequences.[56] Moreover, since Aristotle, moral responsibility accounts in the Western tradition have concentrated primarily upon the *individual's* responsibility. However, as described above, there are scenarios where several agents are involved, making it difficult to identify one single agent as the responsible one. Shared responsibility involves multiple actors contributing to a single outcome, in situations where we cannot determine an individual's causal contributions, and where responsibility is distributed to each actor separately, rather than collectively.[57] Shared responsibility has figured

54. MODEL PENAL CODE § 2.01 (AM. L. INST., Proposed Official Draft 1962) [hereinafter Model Penal Code].

55. Deborah G. Johnson, *Technology with No Human Responsibility?*, 127 J. BUS. ETHICS 707, 713 (2014); *See also* Hughes, *supra* note 43.

56. *See* Fischer, *supra* note 33.

57. André Nollkaemper & Dov Jacobs, *Shared Responsibility in International Law: A Conceptual Framework*, 34 MICH. J. INT'L L. 359, 364 (2013) (discussing shared responsibility in court systems with multiple actors). *See also* Van de Poel et al., *supra* note 40 (discussing collective responsibility).

only minimally in the ethical and legal literature (predominantly in the discussion about states and organizations).[58]

While shared responsibility might be helpful in scenarios where causation does not provide an adequate basis for responsibility, several questions arise. If several persons share responsibility for what happens as a result of what they have done, what factors affect the degree to which each person involved is responsible for the outcome? While several individuals might be involved in the outcome, it is not clear that each have equal degrees of responsibility. Thinking about brain implants, can human entities and non-human entities share responsibility? If yes, what factors affect the degree to which each agent (human or non-human) is responsible for the outcome? For example, in the scenario of a brain implant with a malfunction, it might be easy to establish which agents bear the most responsibility. But in the scenario of an autonomous brain implant, it might not be that easy to figure out if the individual or the implant had a bigger role in the final state of affairs.

Just as the normative foundations of individual moral responsibility remain unsettled, so is the case for shared responsibility. Some authors argue that shared responsibility is based on having shared goals, while for others it is about sharing benefits.[59] Based on the latter, it could be argued that people with brain implants also bear some responsibility since they have employed the implant for their benefit. Moreover, we can argue that engineers, device manufacturers, physicians, in as much as they know the potential risks and enjoy the benefits from selling and implanting these devices, also share moral and legal responsibility in cases of harm from these implants.

C. Agency

The concept of agency is "associated with the idea of being capable of doing something that counts as an act or action,"[60] and being able to think about those actions (that is, being able to translate desires into intentions by combining them with beliefs). There can be different types

58. *Id.*

59. Chinen, *supra* note 27.

60. Kenneth, E. Himma, *Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to be a Moral Agent?*, 11 ETHICS & INFO. TECH. 19, 19 (2008).

of agency (e.g., rational agency);[61] here we are concerned with moral agency.

The concept of moral agency is "ultimately a normative notion that is concerned with the class of beings whose behavior is subject to moral requirements."[62] To be a moral agent involves the capacity to be morally accountable for one's actions. Thus, moral agency requires the capacity to freely choose one's acts and the capacity to be rational. For moral agency, this second capacity is particularly related to the capacity to engage in moral reasoning.[63]

Agency is philosophically intertwined with autonomy. Moreover, autonomy and moral agency go hand in hand with responsibility. When an agent acts autonomously, "it is not possible to hold anyone else responsible for its actions."[64] This link between agency and responsibility is explicitly expressed by the USA Model Penal Code (MPC), Section 2.01, which states that "(1) a person is not guilty of an offense unless his liability is based on conduct that includes a voluntary act or the omission to perform an act of which he is physically capable."[65]

Two other major views of agency are the realist and the attributivist view. According to the former, "the intention to act and the will that leads to the performance of an act are taken to be "real" things, which exist independently of human experience."[66] Whereas on the attributivist view, "intention and free will are attributed, or ascribed, to human agents."[67] This view is supported by the fact that we experience intentions both in our own acts as well as in the acts of others. According to Hage, because attribution is mind-dependent, at least under this view, "agency and responsibility may be attributed to anything," [68] not only humans.

61. For an overview of different conceptions, theories and kinds of agency see Markus Schlosser, *Agency*, STAN. ENCYCLOPEDIA PHIL. (2019), https://plato.stanford.edu/archives/win2019/entries/agency.

62. Himma, *supra* note 60, at 21.

63. *Id.* at 24. Moral reasoning requires a "minimally adequate understanding of moral concepts," the ability to grasp basic moral principles, and "the ability to identify the facts that make one rule relevant and another irrelevant." *Id.*

64. Chinen, *supra* note 27, at 361.

65. Model Penal Code, *supra* note 54. *See also* Herbert Wechsler, *Codification of Criminal Law in the United States: The Model Penal Code*, 68 COLUM. L. REV. 1425, 1436–39 (1968).

66. Jaap Hage, *Theoretical Foundations for the Responsibility of Autonomous Agents*, 25 ARTIFICIAL INTELLIGENCE L. 255, 259 (2017).

67. *Id*. at 260.

68. *Id.* at 261.

D. Revisiting the concept of moral responsibility

The use of brain implants, as with the introduction of other new technologies, have made our practices for holding an actor responsible subject to being continuously challenged and negotiated.[69] In the case of brain implants, we can raise the question, how exactly does a brain implant interfere with agency? Does it interfere with the mental components or the physical ones? Is the integration of the brain implant into our own biology such that the implant integrates with the implicit, subconscious aspects of brain activity and so bypasses the traditional mental capacities of agency? Is the brain implant by itself undermining individual autonomy and agency as it detaches the intention-action causal link, putting into question the voluntary character of the user's actions?

Thinking of the scenario of Mr. B, we would suggest that he is not responsible for his Johnny Cash-loving behavior. The only complication here is that once he knows that the machine has this effect on his agency—and he retains the ability to switch the machine on and off—one might be inclined to argue that he acquires responsibility for those behaviors through his continued use of the device. But this argument should be treated with caution. If the patient needs the device to treat some disabling mental or physical condition, then he is faced with a very stark choice. Indeed, one could argue that patients facing such a stark choice are in a way coerced by the benefits of the implant.

Given the challenges that are posed to our traditional frameworks for dealing with moral responsibility, some have suggested rethinking how and to whom (or to what) moral responsibility is assigned.[70]

E. Can brain implants be moral agents?

Moral responsibility is generally attributed to moral agents, which at least in Western philosophical traditions are exclusively human beings (as they have been regarded as the only beings capable of acting "intentionally and on the basis of free will").[71] However, with advances

69. *See generally* Helen Nissenbaum, *Accountability in a Computerized Society*, 2 SCI. & ENGINEERING ETHICS 25 (1996).

70. Waelbers, *supra* note 41; Peter Kroes & Peter-Paul Verbeek, *Introduction in* 17 THE MORAL STATUS OF TECHNICAL ARTEFACTS 6 (Peter Kroes & Peter-Paul Verbeek eds., 2014); Matthias, *supra* note 27.

71. Hage, *supra* note 66, at 258.

in neuroscience, computer technology, and artificial intelligence systems, who or what qualifies as a moral agent has been questioned.[72]

Some authors argue that as systems become increasingly more autonomous, humans will not be responsible for their behavior. Andreas Matthias characterizes this as a responsibility gap: "[T]here is an increasing class of machine actions, where the traditional ways of responsibility ascription are not compatible with our sense of justice and the moral framework of society because nobody has enough *control* over the machine's actions to be able to assume the responsibility for them."[73] Others like Robert Sparrow argue that "it will no longer be possible [to] hold the programmers/designers responsible for outcomes that they could neither control nor predict. The connection between the programmer/designers and the results of the system, which would ground the attribution of responsibility, is broken by the autonomy of the system."[74]

Other authors reject the responsibility gap by rejecting the "control requirement condition", which states that "a person is responsible for x *only if* the person has control over x."[75] However, as Santoro and colleagues argue, the control requirement is not needed to hold people responsible, as there are cases in which humans are held responsible for outcomes that are outside their control, such as in strict liability cases. Others, such as Nagenborg and colleagues argue that engineers developing and programming brain implants, are responsible for the behavior of the implants they create, on grounds of professional responsibility, even if they cannot control the behavior of their creations.[76]

Some scholars would call this type of responsibility "positive," which emphasizes "the virtue of having or being obliged to have regard for the consequences of his or her actions on others."[77] However, to what extent can engineers, developers, and even physicians be expected to exert themselves to anticipate or prevent the consequences of the (mis)use of their technologies or the technologies they programmed? A more fruitful

72. William Bechtel, *Attributing Responsibility to Computer Systems*, 16 METAPHILOSOPHY 296 (1985); Lucas D. Introna, *Towards a Post-Human Intra-Actional Account of Sociomaterial Agency (and Morality)*, 17 PHIL. ENGINEERING & TECH. 31, 33–34.

73. Matthias, *supra* note 27, at 177.

74. Robert Sparrow, *Killer Robots*, 24 J. APPLIED PHIL. 62, 70 (2007).

75. Matteo Santoro, Dante Marino, & Guglielmo Tamburrini, *Learning Robots Interacting with Humans: From Epistemic Risk to Responsibility*, 22 AI & SOC'Y. 301, 309 (2008).

76. Michael Nagenborg, Rafeal Capurro, Jutta Weber & Christoph Pingel, *Ethical Regulations on Robotics in Europe*, 22 AI & SOC'Y. 349 (2007).

77. Donal Gotterbarn, *Informatics and Professional Responsibility*, 7 SCI. & ENGINEERING ETHICS 221, 227 (2001).

way to look at this is to approach responsibility as having a social function, as what counts is the "social consequences it produces,"[78] including the expectations among it creates among members of society and the incentives it promotes to correct or encourage certain behavior.

Others have responded to the responsibility gap by entertaining the possibility that artificial agents could themselves be held responsible. Certainly, not all brain implants would qualify as moral agents. According to Daniel Dennett, for a system to be a moral agent, is not enough to be an intentional system, but the system itself must be a "higher-order intentional system."[79] That is a system whose behavior relates to its mental states, "capable of framing beliefs about its own beliefs, desires about its desires . . . and so on."[80] Thus, while our examples above might not be instances of moral agents, the development of brain implants with higher-order intentionality remains a real possibility.

In the future, if artificial intelligence-embedded closed loop brain systems get to a point where they can be considered "higher-order intentional systems," it might be reasonable to hold these systems responsible, and not only their makers, or users.[81] The "advanced learning capability will not only make it harder to blame developers and users . . . but will also make it more reasonable to assign responsibility to the [implants],"[82] as humans will be more inclined to treat these agents as if these were responsible for their own behavior.[83] Others, like Asaro, think that like corporations, intelligent agents themselves can bear responsibility.[84] In this case, these entities acquire the status of moral agents.[85]

There have been several authors who oppose the view that computational systems can be regarded as moral agents. Some argue that for a moral agent to be held responsible it has to "be capable of

78. Bernd Carsten Stahl, *Responsible Computers? A Case for Ascribing Quasi-Responsibility to Computers Independent of Personhood or Agency*, 8 ETHICS & INFO. TECH. 205, 210 (2006).

79. Daniel C. Dennett, *When HAL Kills, Who's to Blame? Computer Ethics in* HAL'S LEGACY: 2001'S COMPUTER AS DREAM AND REALITY 351, 354 (David G. Stork ed., 1997).

80. *Id.*

81. Hage, *supra* note 66, at 255.

82. Thomas Hellström, *On the Moral Responsibility of Military Robots*, 15 ETHICS & INFO. TECH. 99, 105 (2012).

83. *Id.*

84. Peter Asaro, *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, *in* ROBOT ETHICS THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS 169 (Patrick Lin, Keith Abney & George A. Bekey eds., 2012).

85. John P. Sullins, *When is a Robot a Moral Agent?*, 6 INT'L REV. INFO. ETHICS 23 (2006) (arguing that in cases where a robot, or other intelligent artifact, fulfills criteria to be a moral agent, it also inherits moral rights as well as responsibilities).

suffering."[86] This particular view of moral agency presupposes that these entities will have to be at least reactive to be held responsible. Others argue that moral agents need to be capable of understanding the meaning of the information that they process.[87]

Floridi and Sanders, in a move around this problem, propose to extend "the class of moral agents" to include artificial agents,[88] disconnecting moral agency from the notion of moral responsibility. That is, artificial agents should be acknowledged as moral agents that can be held accountable, but not responsible.[89] In this case, according to Floridi and Sanders, we can deal directly with the artificial agent for its "bad behavior," for example modifying or deleting it.[90]

A counterargument to Floridi's and Sanders's view is that this would draw attention away from the humans that create, deploy, and use these artificial agents. Johnson, for example, argues that these technologies, although part of the moral world, remain connected to the particular values of their creators and users, as "their functionality has been intentionally created."[91] However, this approach might not adequately capture the type of "higher-order intentional systems" suggested by Dennett. While it is true that traditional brain implants are designed, developed, tested, implanted, initiated, and provided with input, not all new brain implants are provided with particular instructions to perform specified tasks; rather some machine learning-based brain implants are given the "freedom" to find the solution that would lead to the desired outcome. In a way, how different is that from us humans? What is true, however, is that regardless of whether we consider or not brain implants as moral agents, we should not disregard the "social, temporal, cultural, economic or political factors" that shape the development and use of such technologies.[92]

86. *E.g.*, Sparrow, *supra* note 74, at 72; *See also* Asaro, *supra* note 84.

87. Bernd Carsten Stahl, *Information, Ethics, and Computers: The Problem of Autonomous Moral Agents*, 14 MINDS & MACHINES 67 (2004).

88. Luciano Floridi & J.W. Sanders, *On the Morality of Artificial Agents*, 14 MINDS AND MACHINES 349, 361 (2004).

89. *Id.* Floridi and Sanders's view is counter to generally accepted differences between accountability and responsibility, in that being accountable is taken to mean not only being responsible for an action but also ultimately being answerable for that action. But it addresses the problem that while responsibility can be shared, accountability cannot.

90. *Id*.

91. Deborah G. Johnson, *Computer systems: Moral entities but not moral agents*, 8 ETHICS & INFORMATION TECHNOLOGY 195, 201 (2006).

92. MEREL ELISABETH NOORMAN, MIND THE GAP: A CRITIQUE OF HUMAN/TECHNOLOGY ANALOGIES IN ARTIFICIAL AGENTS DISCOURSE 6–7 (2009).

F. Is responsibility distributed among humans and technologies?

Even those that do not see technologies as moral agents still see the active role technology has in shaping human action as part of understanding moral responsibility.[93] In this regard, Johnson argues for a type of shared moral responsibility where the engineer, the brain implant, and the user would all be part of the moral evaluation.[94] Floridi has explored a similar approach with his concept of "distributed moral action,"[95] in which multi-agent systems (including humans along with artificial and hybrid systems) can be held responsible for distributing "morally loaded actions."[96]

Others, rather than focusing on discussing whether artificial agents are moral agents, or whether they can be held responsible, have focused on the role that technology design and development have in addressing responsibility issues.[97] For example, according to Moor, artificial agents can be "implicit ethical agents," as they embody the ethics that their developers inscribed in their design.[98] They can be "explicit ethical agents," representing ethics explicitly and operating on the basis of that knowledge, or "full ethical agents" making ethical judgments that can be justified, much as human beings can.[99] Moor, together with several others, remains skeptical that brain implants, even closed-loop implants designed with machine learning, will ever be full ethical agents. However, the distinction he provides between implicit and explicit ethical agents is helpful for those working on how to design technologies that can be more ethical. Similarly, Johnson and Noorman argue that we should focus on "how to develop artificial agents to ensure that humans can be responsible for their behavior." [100]

93. *See generally*, HANS JONAS, THE IMPERATIVE OF RESPONSIBILITY (Hans Jonas trans., 1984) (1979); Verbeek, *supra* note 49.

94. Johnson, *supra* note 91, at 204.

95. Luciano Floridi, *Distributed Morality in an Information Society*, 19 SCI. & ENGINEERING ETHICS 727 (2012).

96. Luciano Floridi, *Faultless Responsibility: On the Nature and Allocation of Moral Responsibility for Distributed Moral Actions*, 374 PHIL. TRANSACTIONS ROYAL SOC'Y 2083, 2083 (2016). *See also* Floridi, *supra* note 40.

97. Deborah G. Johnson & Merel Noorman, *Recommendations for Future Development of Artificial Agents [Commentary]*, 33 IEEE TECH. & SOC'Y MAG. 22, 24 (2014).

98. James H. Moor, *The Nature, Importance, and Difficulty of Machine Ethics*, IEEE INTELLIGENT SYSTEMS, July/August 2006, at 18, 19.

99. *Id.* at 20.

100. Johnson and Noorman, *supra* note 97, at 22. One of the suggestions they put forward as a way to ensure that humans remain responsible for the behavior of artificial agents is to think about responsibility as a "set of practices," that is, the "established ways (e.g, regulations, medical

Regardless of whether we agree that brain implants can be moral agents, have shared or distributed responsibility with humans, it is clear that with the introduction of even more intelligent technologies, traditional views on moral responsibility are not suitable to capture the complexities of the relationship between responsibility and technology.

It is upon these theories of moral responsibility, agency, and causation that legal liability rests.

IV. LEGAL CONSIDERATIONS

Legal liability for actions that society frowns upon is generally divided into two systems—public and private. Of those systems, we will discuss various issues in criminal and tort liability as ways in which a morally responsible actor could be held liable as a consequence of his actions.[101]

A. *Criminal Law*

Law students are commonly taught that criminal liability is based on the paired concepts of *actus reus*, a guilty act, and *mens rea*, a guilty mind.[102] These common law concepts of actus reus and mens rea have deep roots in Anglo-American law as these concepts look towards the mind of the defendant to gauge whether he is guilty of a crime and to what level he may be guilty.[103] The terms, if not the ideas behind them, are often traced back as far as the 17th century to the English jurist Sir Edward Coke, who penned the legal maxim "*actus non facit reum, nisi mens sit rea*" (an act does not make a person guilty unless his mind is also guilty).[104] The United States Supreme Court supported this maxim in

guidelines, social norms) in which groups and individuals in a community understand, evaluate, and distribute responsibility." *Id.* at 26.

101. This distinction, of course, is a generalization which ignores causes of action and remedies such as administrative penalties.

102. *See, e.g.*, WAYNE R. LAFAVE, 1 SUBSTANTIVE CRIMINAL LAW, § 5.1 (3d. ed. 2019); *see also* Erica Beecher-Monas & Edgar Garcia-Rill, *Actus Reus, Mens Rea and Brain Science: What Do Volition and Intent Really Mean?*, 106 KY. L.J. 265, 267 (2017); Melissa Hamilton, *Reinvigorating Actus Reus: The Case for Involuntary Actions by Veterans with Post-Traumatic Stress Disorder*, 16 BERKELEY J. CRIM. L. 340, 340 (2011); Ian P. Farrell & Justin F. Marceau, *Taking Voluntariness Seriously*, 54 B.C. L. REV. 1545, 1545 (2013).

103. For an in-depth history of intent in criminal law, see Martin R. Gardner, *The Mens Rea Enigma: Observations on the Role of Motive in the Criminal Law Past and Present*, 1993 UTAH L. REV. 635 (1993); Francis Bowes Sayre, *Mens Rea*, 45 HARV. L. REV. 974 (1932).

104. EDWARD COKE, THE THIRD PART OF THE INSTITUTES OF THE LAWS OF ENGLAND, *107. *See also*, Eugene J. Chesney, *Concept of Mens Rea in the Criminal Law*, 29 J. CRIM. L. & CRIMINOLOGY 627, 632 (1939); Hamilton, *supra* note 102, at 343; Gardner, *supra* note 103, at 636.

Morissette v. United States, stating that crime is "a compound concept, generally constituted only from the concurrence of an evil-meaning mind with an evil-doing hand . . . ."[105]

Criminal responsibility, therefore, generally requires a criminal act (actus reus) and a concurrent criminal intent (mens rea).[106] This requirement of both act and intent remains a core of much of modern criminal law today. Generally, both actus reus and mens rea must be met in some connected way for a person to be criminally liable. Actus reus is often thought of as the more external, factual, or objective element, and is a voluntary act or omission that is in some way prohibited.[107] Mens rea is generally considered the more internal, mental, or subjective element and refers to the state of mind of the actor.[108]

1. Actus Reus—A Voluntary Act

A prohibited act is generally considered a necessary element for criminal culpability—otherwise simple thought could be a crime.[109] Indeed, the Model Penal Code specifically states that an individual "is not guilty of an offense unless his liability is based on conduct that includes a voluntary act . . . ."[110] While the act might be thought as the more straightforward of the two elements of mens rea and actus reus, it instead has a history of inconsistency and confusion in scholarly literature and judicial treatment.[111]

But see Farrell & Marceau, *supra* note 102, at 1551 (commenting that the term *actus reus* may actually be a product of the 20th century).

105. Morissette v. United States, 342 U.S. 246, 251 (1952).

106. *See* Fowler v. Padget, 7 Term Rep. 509, 514 (1798) ("It is a principle of natural justice, and of our law, that *actus facit reum nisi mens sit rea*. The intent and the Act must both concur to constitute the crime."); *see also* Gardner, *supra* note 103, at 636; Sayre, *supra* note 103, at 974. However, some specific statutes, most notably those of possession of controlled substances, may not have a true act requirement, and "strict liability" statutes require no *mens rea*. *See, e.g.*, Douglas Husak, *Rethinking the Act Requirement*, 28 CARDOZO L. REV. 2437, 2439–45 (2007); Toke, *infra* note 151 and associated text.

107. Francesca Lagioia & Giovanni Sartor, *AI Systems Under Criminal Law: A Legal Analysis and a Regulatory Perspective*, 33 PHIL. & TECH. 433, 439–40 (2019); Beecher-Monas, *supra* note 102, at 267; LAFAVE, *supra* note 102, § 5.1; Farrell & Marceau, *supra* note 102, at 1549–50. Generally, an omission to act would only meet the requirements for *actus reus* if there existed a duty for the defendant to act in some way.

108. *See* Gardner, *supra* note 103, at 637; LAFAVE, *supra* note 102, § 5.1.

109. *See* Deborah W. Denno, *Crime and Conciousness: Science and Involuntary Acts*, 87 MINN. L. REV. 269, 282–84 (2002); Farrell & Marceau, *supra* note 102, at 1552–54; LAFAVE, *supra* note 102, § 6.1(b). *But see* Husak, *supra* note 106, at 2439–45; Michael Corrado, *Is there an Act Requirement in the Criminal Law?*, 142 U. PA. L. REV. 1529 (1994).

110. MODEL PENAL CODE § 2.01 (AM. LAW INST. 1962).

111. *See, e.g.*, Kevin W. Saunders, *Voluntary Acts and Criminal Law: Justifying Culpability Based on the Existence of Volition*, 49 U. PITT. L. REV. 443, 449 (1988) ("General consensus is

This inconsistency in definition and concept is compounded by the idea that an act must be a voluntary one; after all, it would serve little purpose, either as retribution or deterrence, to punish a person for an act that was not done voluntarily.[112] By requiring a voluntary action before criminal liability can attach, actus reus plays a dual role, preventing the punishment of either "thought crimes" or actions fully outside the control of the defendant.[113] Layering the potential for brain implants to inhibit voluntary actions or incite involuntary actions into the confusion of actus reus may make this element of criminal liability the most difficult for the law to apply.

Although it is generally considered a core concept in criminal law, there is little consensus on what "voluntary act" means. Even the seemingly simple term "act" has divergent treatments in the literature.[114] Sir John William Salmond broadly defined an act as incorporating a bodily movement or action as well as the circumstances surrounding that action and the consequences of that action.[115] In contrast, Oliver Wendell Holmes considered mere body movement, without reference to any surrounding circumstances or consequences, to be an act.[116] Holmes wrote, "An act is always a voluntary muscular contraction, and nothing else."[117] This definition, of course, brings in the problematic term "voluntary."

Defining volition, or the voluntary nature of an act, has been even more contentious than defining an act. Volition has been variously described as a physical manifestation of a defendant's will and determination, behavior that is under the control of the defendant, or

lacking regarding what constitutes an act."); Farrell & Marceau, *supra* note 102, at 1551 ("Although the voluntary act requirement is routinely referred to as an equally 'fundamental principle' of criminal law, there is remarkably little consensus about what it means, why it is required, and how it relates to other elements of criminal liability, including mens rea."); Hamilton, *supra* note 102, at 344–52 (describing various approaches to *actus reus* in the literature and case law).

112. LaFave, *supra* note 102, § 6.1(c).

113. Farrell & Marceau, *supra* note 102, at 1554.

114. *See id.* at 1548–55; Hamilton, *supra* note 102, at 344–48; Saunders, *supra* note 111, at 449–53; LAFAVE, *supra* note 102, § 6.1.

115. JOHN WILLIAM SALMOND, JURISPRUDENCE 370 (10th ed. 1947). For an in-depth discussion of Salmond's concept of an act, see Saunders, *supra* note 111, at 450–53.

116. Corrado, *supra* note 109, at 1529 n.1. Corrado's article contains an extensive discussion of voluntariness and movement as an act. *Id. See also* LAFAVE, *supra* note 102, § 6.1(a).

117. OLIVER WENDELL HOLMES, JR., THE COMMON LAW 91 (1881). Like Holmes's definition, the Model Penal Code defines an act by bodily movement: "'act' or 'action' means a bodily movement whether voluntary or involuntary." MODEL PENAL CODE § 1.13 (AM. LAW INST 1962). The Model Penal Code then requires that an act be voluntary for criminal liability to be imposed. *Id.* § 2.01(1).

simply been identified by what it is not.[118] The Model Penal Code uses this last method. Rather than defining a voluntary act, the Model Penal Code specifically excludes "(a) a reflex or convulsion; (b) a bodily movement during unconsciousness or sleep; (c) conduct during hypnosis or resulting from hypnotic suggestion; (d) a bodily movement that otherwise is not a product of the effort or determination of the actor, either conscious or habitual."[119]

The importance of the Model Penal Code's incorporation of a voluntary act as an element of a crime, and thus the actus reus concept, cannot be overstated. The majority of U.S. states have incorporated the Model Penal Code (in whole or part) into their statutes.[120] Furthermore, most crimes today are statutory, rather than common law.[121]

Michigan, for instance, states in its Revised Statutes that an individual's "criminal liability is based on conduct that includes either a voluntary act or an omission to perform an act or duty that the person is capable of performing."[122] While Michigan does not seem to have supplied a statutory definition for a voluntary act, the statute clearly considers a voluntary act or omission as an element required for a defendant to be guilty of a crime.

Given the debates surrounding actus reus in the literature and the confusion over whether the 'objective' element has a mental component, it should come as no surprise that judicial treatment of actus reus is also inconsistent.[123] These debates are not inconsequential. The interpretation of the voluntariness of an act could impact liability for crimes that have a strict liability mens rea.[124] Similarly, voluntariness as an element of a crime must be proven beyond a reasonable doubt, but lack of voluntariness as a defense moves the burden onto the defendant.[125]

Criminal liability for DBS patients further complicates this picture. Assuming from the ethical discussion above that we grant moral agency

118. *See* LAFAVE, *supra* note 102, § 6.1(c); Hamilton, *supra* note 102, at 344–48; Corrado, *supra* note 109; Husak, *supra* note 106. For an in-depth discussion of the voluntariness aspect of an act in criminal law, *see* Farrell & Marceau, *supra* note 102.

119. MODEL PENAL CODE § 2.01 (AM. LAW INST. 1962).

120. Paul H. Robinson & Markus D. Dubber, *The American Model Penal Code: A Brief Overview*, 10 NEW CRIM. L. REV. 319, 326 (2007).

121. *Id.* at 320.

122. MICH. COMP. LAWS § 8.9(1)(a) (2015) (emphasis added).

123. For extensive discussions of the judicial application of the voluntariness requirement to *actus reus*, *see* Denno, *supra* note 109, at 275–86; Beecher-Monas & Garcia-Rill, *supra* note 102, at 293–96; Hamilton, *supra* note 102, at 348–52; Farrell & Marceau, *supra* note 102, at 1555–67.

124. Farrell & Marceau, *supra* note 102, at 1547.

125. Beecher-Monas & Garcia-Rill, *supra* note 102, at 293–94; Hamilton, *supra* note 102, at 349–51; Farrell & Marceau, *supra* note 102, at 1556–64.

to the brain implant, is the action a voluntary activity of the patient or the brain implant itself? Even if the brain implant has no moral agency, is the voluntary act of getting an implant that influences actions enough to overcome the volition requirement for an unrelated criminal action? Or should the effects of the brain implant be considered along the same lines as a muscle spasm or convulsion? Due to the differences in the application of the voluntariness standard in mens rea, it is impossible to predict how courts will analyze the situation. Additionally, the need for proof as to the impact of the brain implant on volition will turn on the role of voluntariness as an element of the crime or as a defense. Finally, the specific technology at issue, such as the use of the brain implant and the potential use of an AI in conjunction with the brain implant, may also impact the voluntariness analysis.

Despite descriptions of actus reus as comprising the external or objective element of a crime, the inclusion of volition in the act element clearly adds an internal or mental aspect. This has led some commentators and courts to incorrectly conflate the subjective analysis of voluntariness with that of intent in the mens rea element of the crime.[126] It is possible that meeting the intent requirement of mens rea would also meet the voluntariness requirement of actus reus, but the reverse may not be true. For example, a willful, voluntary movement of a leg does not necessarily imply that the actor intended to trip a passerby. Therefore, we will examine mens rea separately.

2. Mens Rea

Mens rea is the state of mind for the voluntary act. Despite its importance in the area of criminal law, mens rea has long been a source of confusion and argument.[127] Various courts and statutes have attempted to describe requisite mens rea with terms such as intentionally, knowingly, willfully, malice aforethought, and so on.[128] Further confusion is supplied by the common law categories of general intent and specific intent.

126. See Farrell & Marceau, *supra* note 102, at 1564–65; Hamilton, *supra* note 102, at 351.

127. *See, e.g.*, Morissette v. United States, 342 U.S. at 252 ("The unanimity with which they have adhered to the central thought that wrongdoing must be conscious to be criminal is emphasized by the variety, disparity and confusion of their definitions of the requisite but elusive mental element."); see also Sayre, *supra* note 103, at 974 ("[W]hen it comes to attaching a precise meaning to *mens rea*, courts and writers are in hopeless disagreement.").

128. *See Morissette*, 342 U.S. at 252 ("However, courts of various jurisdictions, and for the purposes of different offenses, have devised working formulae, if not scientific ones, for the instruction of juries around such terms as 'felonious intent,' 'criminal intent,' 'malice aforethought,' 'guilty knowledge,' 'fraudulent intent,' 'wilfulness,' 'scienter,' to denote guilty knowledge, or 'mens rea,' to signify an evil purpose or mental culpability.").

General intent is the intent to commit the act in question.[129] Therefore, a crime requiring only general intent does not require that the defendant have any desire to violate the law or have any awareness that the act would produce some specific harm. Crimes of specific intent, on the other hand, require the intent to both commit the act and produce a particular result thereby.[130]

Statutes are often unclear as to whether they require simply general or specific intent. In such cases, it is left to the courts to determine whether the legislature desired general or specific intent. The definition of first-degree murder, for instance, may require that an act was committed with the specific intent that it results in the death of another.[131] In contrast, second-degree manslaughter would simply require that the act was intended, rather than the result being intended, and thus is a general intent crime.[132]

The Model Penal Code attempted to alleviate the confusion surrounding *mens rea*[133] by presenting four levels of responsibility, or "descriptive states of mind,"[134] for crimes requiring *mens rea*.[135] These levels, listed from the least to the most serious, are: negligently, recklessly, knowingly, and purposely.[136]

Of these, knowingly and purposely are most aligned with the common law concept of intent. The Model Penal Code defines a person as acting purposely "(i) if the element [of an offense] involves the nature of his conduct or a result thereof, it is his conscious object to engage in conduct of that nature or to cause such a result; and (ii) if the element involves the attendant circumstances, he is aware of the existence of such

129. *See* Eric A. Johnson, *Understanding General and Specific Intent: Eight Things I Know for Sure*, 13 OHIO ST. J. CRIM. L. 521, 530 (2016).

130. *Id.* at 525.

131. For example, the New York statute for first degree murder states, in pertinent part: "A person is guilty of murder in the first degree when . . . [w]ith intent to cause the death of another person, he causes the death of such person or of a third person. . ." N.Y. PENAL LAW § 125.27(1) (McKinney 2019). The state of Virginia defines first degree murder in pertinent part as "[m]urder, other than capital murder, by poison, lying in wait, imprisonment, starving, or by any willful, deliberate, and premeditated killing, or in the commission of, or attempt to commit, arson, rape, forcible sodomy, inanimate or animate object sexual penetration, robbery, burglary or abduction . . . ." VA. CODE ANN. § 18.2-32 (2020).

132. *See, e.g.*, N.Y. PENAL LAW § 125.15 (McKinney 2019) ("A person is guilty of manslaughter in the second degree when . . . [h]e recklessly causes the death of another person . . . .")

133. Gardner, *supra* note 103, at 682.

134. *See id.* at 688–90.

135. The category of strict liability crimes have no requisite mental component but would still, theoretically, require a voluntary *actus reus*. R.A. Wasserstrom, *Strict Liability in the Criminal Law*, 12 STAN. L. REV. 731, 733 n. 15 (1960).

136. MODEL PENAL CODE § 2.02(1) (AM. LAW INST. 1962).

circumstances or he believes or hopes that they exist."[137] In contrast, the Model Penal Code defines a person's actions as knowingly "(i) if the element [of an offense] involves the nature of his conduct or the attendant circumstances, he is aware that his conduct is of that nature or that such circumstances exist; and (ii) if the element involves a result of his conduct, he is aware that it is practically certain that his conduct will cause such a result."[138]

Therefore, "purposely" requires that it was the "conscious object," or the specific intent, of the person to engage in the criminal act or produce a particular result.[139] "Knowingly," however, is satisfied if the person is aware a particular result is almost certain to occur due to the act taken.[140]

Despite the Model Penal Code's attempt to offer a clarified structure for mens rea analysis, the mental component of criminal law remains a murky and confusing area. Recent technological advances in diverse neurosciences have only produced more challenges for this analysis. Some studies have found that neuroscience evidence has been generally limited to the mitigation of punishment, diminished capacity, and ineffective assistance of counsel.[141] However, while limited, these cases exist across a large swath of criminal charges.[142]

Brain implants introduce the potential of an external stimulus that may alter the intent or produce actions entirely in the absence of a person's intent. For instance, brain implants have the potential to reduce the capacity of a person to possess a requisite mental state. In particular, brain implants may implicate the doctrine of diminished capacity. Diminished capacity is based on the idea that a person cannot be guilty of a crime if she is incapable of forming the necessary intent to be criminally liable.

Diminished capacity, accepted in many jurisdictions, allows the defense to attack the prosecution's evidence of a particular mental state.[143] This evidence may consist of testimony by mental health clinicians to demonstrate that the defendant had a mental disease that would preclude her from forming the requisite mens rea. For example, Ms. Q, who

137. *Id.* § 2.02.

138. *Id.*

139. *Id.* § 2.02(2)(a).

140. *Id.* § 2.02(2)(b).

141. *See, e.g.,* Nita A. Farahany, *Neuroscience and Behavioral Genetics in US Criminal Law: An Empirical Analysis*, 2 J.L. & BIOSCIENCES 485 (2015); Deborah W. Denno, *The Myth of the Double-Edged Sword: An Empirical Study of Neuroscience Evidence in Criminal Cases*, 56 B.C. L. REV. 493 (2015).

142. Henry T. Greely & Nita A. Farahany, *Neuroscience and the Criminal Justice System*, 2 ANN. REV. CRIMINOLOGY 451, 455 (2019).

143. Henry F. Fradella, *From Insanity to Beyond Diminished Capacity: Mental Illness and Criminal Excuse in the Post-*Clark *Era*, 18 U. FLA. J.L. & PUB. POL'Y 7, 47 (2007).

swerved her car into a crowded bus stop, killing several bystanders, might be charged with second-degree murder[144] and manslaughter.[145] If the defense can show evidence that she did not intentionally swerve the car in order to kill people, but that she only intended to swerve, then she may be convicted of manslaughter, but not second-degree murder. Therefore, evidence of the effects of brain implants is particularly relevant in crimes requiring specific intent. Some jurisdictions follow the Model Penal Code approach, allowing diminished capacity evidence in all crimes with a requisite mental state. Other jurisdictions that use the doctrine of diminished capacity restrict it to crimes requiring specific intent, or sometimes even just murder.[146]

Our various fictitious scenarios illustrate some of the problems that brain implants may impart to the *mens rea* analysis. Mr. D's use of DBS would render him in such an incapacitated state that he would need permanent psychiatric care. It is not a stretch to say that he would have diminished capacity while using DBS. On the other hand, in our scenario he *chose* to become incapacitated in order to treat a disease. To what extent should he then bear the responsibility for acts that he commits in such a state? Mr. B's love of Johnny Cash and his corresponding debt is due to DBS. But to what extent would his choice to commit theft be? Without the influence of DBS, he might have chosen to sell his Johnny Cash collection rather than steal. Conversely, even with DBS, he might have had other options to legally deal with his debt—even if it meant losing his house. The DBS would perhaps be influencing his choices but would not be directly urging him to steal. But how would we practically differentiate the two?

Similarly, the strength of the hypothetical Ms. Q's urge to swerve would be difficult to prove. The urge could be anywhere on the spectrum, from uncontrollable muscle spasms to a desire to jerk the wheel and see what happens. How should the difficulty in proving the impact of DBS affect legal liability for the action?

144. "[C]riminal homicide constitutes murder when: (a) it is committed purposely or knowingly; or (b) it is committed recklessly under circumstances manifesting extreme indifference to the value of human life." MPC 210.2.

145. "Criminal homicide constitutes manslaughter when: (a) it is committed recklessly; or (b) a homicide which would otherwise be murder is committed under the influence of extreme mental or emotional disturbance for which there is reasonable explanation or excuse." MPC 210.3.

146. Henry F. Fradella, *From Insanity to Beyond Diminished Capacity: Mental Illness and Criminal Excuse in the Post-*Clark *Era*, 18 U. FLA. J.L. & PUB. POL'Y 7, 47 (2007).

B. Civil Liability

The private system of tort liability allows a person to seek recompense from a second individual in court, even if there is no criminal liability. Tort liability may allow multiple parties to be held responsible for an action, depending on the theory put forth. For instance, in Ms. Q's car accident, an injured party might sue Ms. Q as well as the brain implant manufacturer and the surgeon who installed it. While torts may be characterized in many ways, for our purposes, we will divide torts into those based on product liability, medical malpractice, and patient liability.

1. Product Liability (Strict Liability)

One theory under which product manufacturers have generally been held liable for harms caused by their products is product liability. Product liability is a strict liability tort in which no intent is necessary, thus muting much of our discussion of agency and moral responsibility.[147] However, even here, the peculiarities of DBS and its software-driven (and potentially AI-driven) effects complicate the analysis of liability.

Additionally, the ability to sue medical device manufacturers under a products liability cause of action has been sharply abrogated under statutes and recent case law. Before looking at potential pre-emption issues, a look at general product liability law and FDA regulatory procedures for medical devices is necessary.

Products liability theory imposes strict liability on a product manufacturer if the product sold has defects in its design or manufacture, or in the product warnings.[148] While manufacture defects often involve a straightforward analysis of the product to the manufacturer's own designs, design defects are more difficult to determine. Defective designs can be proven under either a consumer expectations test or a cost-benefit approach. Under the consumer expectations test, "[t]he article sold must be dangerous to an extent beyond that which would be contemplated by the ordinary consumer who purchases it, with the ordinary knowledge to the community as to its characteristics."[149] The cost-benefit approach is used "when the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design . . . and the omission of the alternative design renders the product

147. *See* RESTATEMENT (SECOND) OF TORTS § 402A (AM. LAW INST. 1965).
148. *Id.* cmt. f–k.
149. *Id.* cmt. i.

not reasonably safe . . . ."[150] These tests require the court to determine the standard of proper design.[151]

Moreover, the Restatement (Second) of Torts recognizes that plaintiffs hurt by defective products include consumers, users, and bystanders.[152] Therefore, product liability causes of action generally are open to both DBS patients and those harmed by the patients' actions. However, recent interpretations of the statutes surrounding medical device regulation have abrogated product liability causes of action.

In *Riegel v. Medtronic*,[153] the Supreme Court held that the preemption clause of the Medical Device Amendments of 1976 (MDA)[154] bars state claims challenging safety or effectiveness of a medical device that has been approved by the FDA. Manufacturing defects in which the product does not comply with the designs submitted to the FDA are not covered by this preemption. But the *Riegel* decision limits product liability for design and warning defects for devices that require FDA pre-market approval. Not all medical devices require pre-market approval. Medical devices are broadly defined as:

> an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article . . . which is . . . intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals.[155]

Under this broad definition, medical devices are subdivided into three classes based on risk assessments of the intended use of the product.[156]

Each medical device class has its own regulatory requirements,[157] which in turn impacts the product liability under which manufacturers operate. Class I devices do not introduce substantial safety risks to patients. Class II and class III pose safety risks to patients with increasing severity in each classification. Class I and class II devices do not require

150. RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(b) (AM. LAW INST. 1998).

151. Michael J. Toke, *Restatement (Third) of Torts and Design Defectiveness in American Products Liability Law*, 5 CORNELL J.L. & PUB. POL'Y 239, 241 (1996).

152. *See* RESTATEMENT (SECOND) OF TORTS § 402A caveat 1 (AM. LAW INST. 1965).

153. Riegel v. Medtronic, 552 U.S. 312 (2008).

154. 21 U.S.C. § 360k(a) (2018).

155. 21 U.S.C. § 321(h) (2012).

156. *Device Classification Panels*, U.S. FOOD & DRUG ADMIN., U.S. FOOD & DRUG ADMIN. (Aug. 31, 2018), https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm051530.htm [https://perma.cc/U4KM-QU53].

157. For an overview of the medical device review process, see Ashwani Sastry, *Overview of the US FDA Medical Device Approval Process*, 16 CURRENT CARDIOLOGY REP. 494 (2014).

pre-market approval, but rather a pre-market notification (often called a 501k notification after the authorizing statute). Class III devices require submission of clinical trial data[158] and must obtain pre-market approval (PMA) from the FDA before sales.[159] Therefore, under the holding of *Riegel v. Medtronic*, only class III medical devices are generally exempt from product liability claims based on the safety or effectiveness of the device.

Due to its implantation into, and effects on, the brain, DBS is accomplished using class III, FDA-approved medical devices.[160] Therefore, while product liability claims will continue to be available for manufacturing defects of DBS devices, it is unlikely that products liability will extend to problems arising from the initial design. However, there are two theories under which DBS devices may still be subject to product liability causes of action.

First, while the brain implanted portion of the DBS device is likely a class III medical device, the FDA may not review the software portion of the product at the same level.[161] The application of product liability pre-emption in such cases has been called into question.[162] In *Shuker v. Smith & Nephew*, the Third Circuit opened the door to analyzing each component of a medical device separately for preemption purposes.[163] Since the algorithms of the software are often a "black box" for the FDA regulators,[164] it is possible that product liability claims for failure of AI systems could be divorced from the pre-market approval preemption under *Riegel*.

Second, DBS devices may evolve over time *within the implanted device itself*. For instance, the cyber security of a device may be updated regularly, and an AI that learns from prior rounds of stimulation and response may reprogram itself to make different decisions in the future. While the initial software will have been approved by the FDA, should

158. *Id.* at 496.

159. 21 C.F.R. § 860.3(c)(1)–(3).

160. Class III medical devices "sustain or support life, are implanted, or present potential unreasonable risk of illness or injury. [Other] [e]xamples of Class III devices include implantable pacemakers and breast implants." *Learn if a Medical Device Has Been Cleared by FDA for Marketing*, U.S. FOOD & DRUG ADMIN. (Dec. 29, 2017), https://www.fda.gov/medical-devices/consumers-medical-devices/learn-if-medical-device-has-been-cleared-fda-marketing [https://perma.cc/RBN3-25N5].

161. *See* W. Nicholson Price II, *Regulating Black-Box Medicine*, 116 MICH. L. REV. 421, 452 (2017).

162. Charlotte Tschider, *Preempting the Artificially Intelligent Machine*, BYU L.R. (forthcoming), https://ssrn.com/abstract=3443987 [https://perma.cc/TVK8-D3TA].

163. Shuker v. Smith & Nephew, PLC, 885 F.3d 760, 770–74 (3d Cir. 2018).

164. *See* W. Nicholson Price II, *Black-Box Medicine*, 28 HARV. J.L. & TECH. 419 (2015).

the pre-emption that follows from that approval extend to changes made as the AI evolves? Additionally, as the AI evolves, it becomes less tangibly linked to the device developer in the first instance. Returning to the discussion of moral responsibility, can an AI be held liable separate from its programmer? What would such liability look like—an AI does not have its own bank account. If an AI is liable in some moral fashion, does that separate liability shield the programmer from liability for creating the AI that caused harm? The arguments of Nagenborg and colleagues[165] would suggest professional responsibility should continue to impart liability to the engineers designing and programing brain implants, even if they cannot control them.

Finally, another potential exception to the design defect may occur in the scenario of cyber hacking. Cybersecurity of medical devices is constantly evolving to address new cyber threats. While the initial cybersecurity is part of the FDA requirements when needed, and the FDA encourages frequent updating of cybersecurity, those updates do not need to be approved by the FDA before implementation.[166] Therefore, a DBS device manufacturer may not be preempted by the MDA bar against tort actions under *Riegel*.

One way to avoid the problem of the lack of liability for harm caused by the DBS device regulation in the absence of strict liability for AI might be to regulate the DBS AI as a medical provider.[167] Many medical devices already act more as diagnosticians than mere devices.[168] Such an option leads to an analysis of the second group of torts based on negligence.

2. Professional Negligence

State-imposed duties of care generally define when negligence liability lies. Negligence arises when a duty of care is breached and that breach causes an injury. As an example of negligence, a breach of the doctor-patient duty of care could give rise to an action for malpractice, aka, professional negligence. There are two potential avenues to pursue

165. Michael Nagenborg, Rafeal Capurro, Jutta Weber & Christoph Pingel, *Ethical Regulations on Robotics in Europe*, 22 AI & SOC'Y. 349 (2007).

166. *The FDA's Role in Medical Device Cybersecurity*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/UCM544684.pdf [https://perma.cc/7VEU-4MTL]

167. *See* Jane R. Bambauer, *Dr. Robot*, 51 U.C. DAVIS L. REV. 383 (2017).

168. *See* Adam Candeub, *Digital Medicine, the FDA, and the First Amendment*, 49 GA. L. REV. 933 (2015) (discussing phone apps such as Dr. Mole, an app that tells customers if their moles are likely to be benign or cancerous).

medical malpractice claims in the DBS context—against the surgeon and against the AI itself.

Medical malpractice requires four elements: fiduciary duty, breach of the duty, causation, and damages.[169] Within a doctor-patient relationship, doctors are held to a standard of care based upon their fiduciary relationship with the patient.[170] The standard is based on the reasonable care a doctor would give[171]—often based on the standard of care set forth by medical professional societies.[172]

As doctors begin to use DBS or other techniques that increasingly rely on AI or that have previously unknown side effects, the learning curve for the new medical device may make the implementation of the device risky enough to open the surgeon to liability.[173] Allowing an AI to make dosing decisions using feedback algorithms may be helpful to the patient, but could impact the level of care the doctor is perceived to be giving unless there is continued oversight similar to that given to other agents.

Alternatively, Jane Bambauer has proposed regulating AI devices as medical professionals subject to tort liability.[174] Such a proposal would be interesting. Regulating an AI program as a medical professional would treat the AI as a principal rather than an agent of the medical doctor.[175] In such a case, the AI's liability would likely be imputed to the medical device manufacturer, and companies that produce DBS controlled by AI would likely seek coverage for malpractice insurance similar to medical professionals today.

3. Patient Tort Liability

In civil liability, we return to the liability of the brain implant patient as we examine negligence and intentional tort liability of the patient. Returning to the hypothetical of Ms. Q, who swerved into the bus shelter, we can begin to see the complexity of her potential tort liability.

---

169. 3 BARRY A. LINDAHL, MODERN TORT LAW: LIABILITY AND LITIGATION § 24:1 (2d ed. 2020).

170. *Id.* § 24:1, 24:3.

171. *Id.* § 24:15.

172. *Id.* § 24:35.

173. *See* Frank Griffin, *The Trouble with the Curve: Manufacturer and Surgeon Liability for "Learning Curves" Associated with Unreliably-Screened Implantable Medical Devices*, 69 ARK. L. REV. 755, 765 (2016).

174. Bambauer, *supra* note 167.

175. Hannah R. Sullivan & Scott J. Schweikart, *Are Current Tort Liability Doctrines Adequate for Addressing Injury Caused by AI?*, 21 A.M.A J. ETHICS 160, 163–65 (2019).

Intentional torts are based on actions that cause reasonably foreseeable harm to an individual.[176] Intentional torts generally include assault,[177] theft (conversion),[178] and fraud.[179] As with criminal law, confusion exists as to how to interpret the intent requirement of these torts.[180] Since intentional torts require an intentional action, the impact of brain implants on the state of mind of the actor should matter as much as described in criminal liability. For instance, debate exists as to whether intent encompasses merely the intent to act (single intent) or intent to perform an action and cause a harm (dual intent).[181]

Although scholars conceded that the debate between single and dual intent currently has few real-world consequences,[182] it could be important in the DBS context. Ms. Q's urge to swerve the car poses an interesting question about intent. It is unclear whether her response was an intentional response to an urge from an outside stimulus or an autonomic muscle response. But even if it were an intentional action, in no case did Ms. Q intend to swerve into a crowd of people. If it is the intent to produce the action that legally matters for intentional torts, she could face liability. Conversely, a dual intent standard would take into account the lack of intent to harm others based on the improper operation of the brain implant.

While negligence does not depend on a state of mind, it does depend on a number of factors that may be impacted by brain implants. Each jurisdiction defines negligence slightly differently, but generally, negligence requires (1) breach of a duty, (2) proximate cause, and (3) damages.[183] Vehicle accidents are often brought as negligence actions.

The duty owed to society by a patient with a brain implant is currently amorphous. Duties arise in society from social custom based on the balance of societal safety against personal freedom.[184] Even if Ms. Q's urge to swerve the car was not intentional, driving while under the

176. "The word "intent" is used throughout the RESTATEMENT of this Subject to denote that the actor desires to cause consequences of his act, or that he believes that the consequences are substantially certain to result from it." RESTATEMENT (SECOND) OF TORTS § 8A (AM. LAW INST. 1965).

177. *See id.* § 21.

178. *See id.* § 222A.

179. *See id.* § 525.

180. For a fuller discussion of the controversy as to intent in tort law, see Nancy J. Moore, *Intent and Consent in the Tort of Battery: Confusion and Controversy*, 61 AM. U.L. REV. 1585 (2012).

181. *Id.* at 1588.

182. *E.g.*, *id.* at 1593.

183. David G. Owen, *The Five Elements of Negligence*, 35 HOFSTRA L. REV. 1671, 1672–74 (2007).

184. *See id.* at 1674–75.

influence of a brain implant could be deemed a breach of a duty of safety by the patient in a similar manner to patients with epileptic seizures.[185]

Perhaps even the medical device company or the software programmers could be reached through this negligence action. Generally, the superseding cause doctrine, which says that "[w]hen a force of nature or an independent act is also a factual cause of harm, an actor's liability is limited to those harms that result from the risks that made the actor's conduct tortious,"[186] would shield them from liability due to the intervening negligent conduct of Ms. Q. However, the superseding cause doctrine could be less helpful to AI developers as courts move to a foreseeability standard if courts believe DBS-related accidents are foreseeable.[187]

V. CONCLUSION

Researchers, clinicians, manufacturers, and regulatory bodies should cooperate to minimize the challenges and risks posed to responsibility frameworks by brain implants. Considering the impact brain implants can have on our moral and legal notions of responsibility, we need to discuss whether and when brain interventions should excuse people. New technologies often require some modification or extension of existing legal mechanisms, so understanding the issues of responsibility raised by brain implants will become increasingly important. It is possible that soon we will start hearing in courtrooms: "It's not my fault. My brain implant made me do it." And so, the question becomes, who will sing the Folsom Prison Blues[188] in this new era of brain implants?

185. For an example of how Ohio treats negligence actions in the context of epilepsy, see Kathryn Kramer, *Shifting and Seizing: A Call to Reform Ohio's Outdated Restrictions on Drivers with Epilepsy*, 22 J.L. & HEALTH 343, 357–58 (2009).

186. RESTATEMENT (THIRD) OF TORTS: PHYSICAL & EMOTIONAL HARM § 34 (AM. LAW INST. 2010).

187. *See* Weston Kowert, *The Foreseeability Human-Artificial Intelligence Interactions*, 96 TEX. L. REV. 181, 189–91 (2017).

188. Johnny Cash, *Folsom Prison Blues*, YOUTUBE, https://www.youtube.com/watch?v=6ZPToXstS8M [https://perma.cc/XKT9-27DN].