ERICKSON SK. Blaming the Brain. MINN. J.L. SCI. & TECH. 2010;11(1):27-77.

Blaming the Brain

Steven K. Erickson*

I. INTRODUCTION

People are more than their brains. Legal and social traditions have long held people accountable for their behavior under the presumption that most behavior is intentional and the product of conscious decision-making.[1] The elementary premise that actions have legal and social consequences is firmly embedded in the vernacular of most cultures.[2] When

* Steven K. Erickson is a visiting Associate Professor of Law, University of Missouri School of Law. He is deeply indebted to the thoughtful comments provided by Stephanos Bibas, Stephen Morse, Jon Klick, Amy Wax, Jeff Rachlinski, David Skeel, Michael Perlin, Paul Litton and the participants of the ad hoc faculty workshop series at the University of Pennsylvania Law School. Bill Draper provided excellent research assistance.

1. *See* Morissette v. United States, 342 U.S. 246, 250 (1952) ("The contention that an injury can amount to a crime only when inflicted with by intention is no provincial or transient notion. It is as universal and persistent in mature systems of law as belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil."); Roscoe Pound, *Introduction* to FRANCIS BOWES SAYRE, A SELECTION OF CASES ON CRIMINAL LAW, at xxxvi-xxxvii (1927) ("Historically, our substantive criminal law is based upon a theory of punishing the vicious will. It postulates a free agent confronted with a choice between doing right and doing wrong and choosing freely to do wrong.").

2. *See* Paul H. Robinson & John M. Darley, *Intuitions of Justice: Implications for Criminal Law and Justice Policy*, 81 S. CAL. L. REV. 1, 8–11

someone breaches a legal or social code, the whole person is considered responsible for her behavior and not some constituent part of her entity that bears the punishment. Unlike machines, people are measured in their totality—a bad part does not render a person worthless.

Yet emerging conceptions of personhood generated by cognitive neuroscience suggest something very different from this entrenched view.[3] Instead of people, cognitive neuroscience posits brains as the exclusive agents of behavior and suggests brains are incapable of blame because of their mechanical and determined nature.[4] Minds and brains are held synonymous—thoughts, desires, and behaviors are regarded as no more than the yield of fixed neuronal tissue.[5] All mentation is considered exclusively the product of brain structure and function, which is accessible, measurable, and predictable using an array of novel technologies understood by a select few. This view has led to speculation that an inevitable and radical overhaul of legal and social constructions of personal responsibility is invariably at hand.[6] Much of the recent legal

---

(2007); James Q. Wilson, *The Moral Sense*, 87 AM. POL. SCI. REV. 1, 1 (1993).

3. *See* Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts*, 87 MINN. L. REV. 269, 271–74 (2002); Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y B 1775, 1780–81 (2004); Maureen Sie & Arno Wouters, *The Real Challenge to Free Will and Responsibility*, 12 TRENDS COGNITIVE SCI. 3, 3 (2007) ("Recent developments in the behavioral-, cognitive- and neurosciences indicate that, more often than not, we act in an automatic and unaware fashion, making up reasons only as we go along.").

4. For an overview of the arguments see Adina Roskies, *Neuroscientific Challenges to Free Will and Responsibility*, 10 TRENDS COGNITIVE SCI. 419 (2006); see also Jeffery M. Schwartz et al., *Quantum Physics in Neuroscience and Psychology: A Neurophsyical Model of Mind-Brain Interaction*, 360 PHIL. TRANSACTIONS ROYAL SOC'Y B 1309, 1313–15 (2005) (discussing ways classic physics lends to understanding mental capacity: (1) as a casually impotent sideshow of the brain or (2) as the very same thing as some pattern of the brain's various tiny parts).

5. As it was famously put "your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than a vast assembly of nerve cells and their associated molecules." FRANCIS CRICK, THE ASTONISHING HYPOTHESIS: THE SCIENTIFIC SEARCH FOR THE SOUL 3 (1994); *see also* MARVIN L. MINSKY, THE SOCIETY OF MIND 287 (1986) ("Minds are simply what brains do."); Greene & Cohen, *supra* note 3, at 1779 ("It is not as if there is *you*, the composer, and then *your brain*, the orchestra. You *are* your brain.").

6. *See* DERK PEREBOOM, LIVING WITHOUT FREE WILL 158–86 (2001); Oliver R. Goodenough & Kristin Prehn, *A Neuroscientific Approach to*

scholarship concerned with criminal responsibility as of late has invested heavily in the notion that the findings of biological sciences promise a fundamental shift away from orthodox notions of criminal liability.[7] Some view the incorporation of neuroscience within the law as a harbinger of change for sentencing practices;[8] others as a welcome aid in shifting punishment away from a retributive-based enterprise to one solely focused on future offender propensities;[9] while still others envision a dramatic diminution of the jury's province.[10] All share the belief that the impact of neuroscience on the law in the coming years will be inevitable, dramatic, and will fundamentally alter the way the law does business.[11] And nowhere is this promise endorsed with more gusto than in discussions of responsibility and criminal liability.

Such renditions of culpability under this rubric of the neuro-person are usually couched with the belief that incorporating cognitive neuroscience into our legal institutions

---

Normative Judgment in Law and Justice, 359 PHIL. TRANSACTIONS ROYAL SOC'Y B 1709, 1718–21 (2004) (describing how neuroscience will change conceptions of morals, justice and normative judgment). *But see* Stephen J. Morse, *Brain and Blame* 84 GEO. L.J. 527, 531 (1996) (describing the "fundamental psycholegal error" of assuming causation equates to excuse for matters of criminal culpability).

7. These views are undoubtedly embedded in a strict biological view of personhood and personal agency that the law has long resisted. *See, e.g.,* Owen D. Jones & Timothy H. Goldsmith, *Law and Behavioral Biology*, 105 COLUM. L. REV. 405, 419 (2005) (decrying "the near-total absence of recognition in legal thinking that all behavior, and all the brain activity that perceives and directs it, are fundamentally biological phenomena . . .").

8. *See* Henry T. Greely, *Neuroscience and Criminal Justice: Not Responsibility but Treatment*, 56 U. KAN. L. REV. 1103, 1103–04 (2008) ("[A]dvances in neuroscience will change, dramatically, the criminal justice system . . . . [W]e may see major changes in how crimes are investigated, in how trials are conducted, in how sentencing decisions are reached, and in what kinds of sentences are imposed.").

9. *See* Christopher Slobogin, *The Civilization of the Criminal Law*, 58 VAND. L. REV. 121, 157–65 (2005) (describing how neuroscience supports a forward-oriented theory of criminal punishment).

10. *See* Julia Seaman, *Black Boxes*, 58 EMORY L.J. 427, 466–75 (2008) (arguing that neuroscience may supplant the juror's role in determining witness credibility and unveil the secrecy of juror decision-making).

11. The pace at which legal scholarship has embraced neurolaw is evident in the numerous professional conferences, institutes, and large-scale projects undertaken at various law schools. *See* Ken Strutin, *Neurolaw and Criminal Justice*, LLRX.COM, Dec. 28, 2008, http://www.llrx.com/features/neurolaw.htm. The title of one such conference held at the Sandra Day O'Connor College of Law was telling: "The Law and Ethics of Brain Scanning: The Next Big Thing Coming Soon to a Courtroom Near You?"

will lead to more just and compassionate outcomes for criminal defendants.[12] The idea that current regimes of culpability and punishment are influenced imprudently by retributive principles of justice looms large in the neurolaw discussion.[13] Proponents of the neuro-person model suggest that blaming people for behavior that is merely a product of their brains is both foolish and unfair.[14] If the brain and the mind are the same and behavior is entirely determined a brain's micro-events, which are largely (if not entirely) automatic and unconscious, then it follows that blame has no place in modern conceptions of agency and criminal law. Indeed, the rise of the neuro-person model coincides with the emergence of the therapeutic justice movement, which seeks to supplant traditional criminal punishment frameworks with behavior-based interventional models of criminal justice under the idea that criminal conduct is essentially a mental health phenomenon.[15] People commit crimes not because they are

12. *See, e.g.*, PAUL M. CHURCHLAND, THE ENGINE OF REASON, THE SEAT OF THE SOUL: A PHILOSOPHICAL JOURNEY INTO THE BRAIN 309 (1996) ("[W]e are likely to see a revolution in the ways that society deals with the broad spectrum of pathological social behavior. A neurally informed and technologically sophisticated society will be able to make judgments reliably and do things effectively, where the current practice is groping and impotent."); Jana L. Bufkin & Vickie R. Luttrel, *Neuroimaging Studies of Aggressive and Violence Behavior: Current Findings and Implications for Criminology and Criminal Justice*, 6 TRAUMA, VIOLENCE, & ABUSE 176, 186 (2005)

> Justice is typically defined as just deserts. As a result, legal variables with no inherent explanatory worth are summoned to justify less-than-stellar community-level interventions and unproductive institutionalization. In contrast, the type of justice conceived in an interdisciplinary program of study is rehabilitation oriented. Its programs and policies are aimed at therapeutic justice.

Id.

13. At the heart of this argument is retribution's focus on blame during the commission of the crime compared to the forward-focused instrumentality crime-control polices ostensibly provided by neuroscience-based conceptions of agency which fix prediction of future harm, intervention, and incapacitation as the preferred goals of criminal justice. *See* Greene & Cohen, *supra* note 3, at 1783–84; Slobogin, *supra* note 9, at 122.

14. *See, e.g.*, Robert M. Sapolsky, *The Frontal Cortex and the Criminal Justice System*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y B 1787, 1794 (2004) ("[A]lthough it may seem dehumanizing to medicalize people into being broken cars, it can be vastly more humane than moralizing them into being sinners.").

15. *See* GERRY JOHNSTONE, RESTORATIVE JUSTICE 94 (2002) ("Proponents of therapeutic interventions regard much criminal behaviour as symptomatic of underlying psychiatric disorder. . . .").

motivated by greed or malice but because they are sick and in need of treatment—they have impaired brains.

But there are problems with the neuro-person model. Despite the claim by neuroscientists that the mind is accessible, measurable, and predictable, there are good reasons for skepticism on all of these fronts. Cognitive neuroscience, like all fields of science, utilizes assumptions to generate its conclusions. Chief among these is that the complexity of the mind can be understood by examining localized areas of the brain which are presumed indicative of how people *think*. The proliferation of various brain imaging technologies has endowed the popular press and public with a powerful and accessible model of the mind, which propagates this localization view of mentation.[16] This model has been augmented by an abundance of highly publicized studies suggesting a link between certain defined behaviors and some part of the brain by way of impressive pictures of measured brain activity. Many of these studies claim to have located areas of the brain responsible for addiction,[17] violence,[18] wisdom,[19] and even morality.[20] These exceedingly complex behaviors are said to originate in highly defined, independent modules of cognitive function located within specific regions of the brain. In a scant twenty years, pictorial brain imaging has become common parlance for understanding all aspects of human thought and

16. One study documented over 130 unique press articles covering brain imaging between 1994–2004—a number that has surely multiplied in succeeding years. *See* Eric Racine et al., *Brain Imaging: A Decade of Coverage in the Print Media*, 28 SCI. COMM. 122, 128 (2006); *see also, e.g.,* Robert Lee Hotz, *The Brain, Your Honor, Will Take the Witness Stand*, WALL ST. J., Jan. 16, 2009, at A7 (reporting on researchers who used a brain scanner to examine how brain cells behave when assessing criminal responsibility and meting out sentences); Jeffery Rosen, *The Brain on the Stand: How Neuroscience is Transforming the Legal System*, N.Y. TIMES MAG., March 11, 2007, at 48 (reporting generally on how recent neuroscience discoveries is transforming the legal system).

17. *See, e.g.*, Rita Z. Goldstein et al., *Subjective Sensitivity to Monetary Gradients is Associated with Frontolimbic Activation to Reward in Cocaine Abusers*, 87 DRUG & ALCOHOL DEPENDENCE 233 *passim* (2007).

18. *See, e.g.*, Andreas Meyer-Lindenberg et al., *Neural Mechanisms of Genetic Risk for Impulsivity and Violence in Humans*, 103 PROC.NAT'L ACAD. SCI. U.S. 6269 *passim* (2006).

19. *See, e.g.*, Thomas W. Meeks & Dilip V. Jeste, *Neurobiology of Wisdom*, 66 ARCHIVES GEN. PSYCHIATRY 355 *passim* (2009).

20. *See, e.g.*, Adrian Raine & Yaling Yang, *Neural Foundations to Moral Reasoning and Antisocial Behavior*, 1 SOC. COGNITIVE & AFFECTIVE NEUROSCIENCE 203 *passim* (2006).

behavior.

Indeed, there appears to be no limit to explaining all aspects of humanity under the neuro-person model–only technology limits its application. Yet it is inescapable that the novel and powerful technology of brain imaging available to neuroscientists invariably drives their conception of the mind.[21] And that technology assumes that brains operate linearly, are stable over time, and that behavior is predictable from localized brain states. Much evidence abounds, however, that brains do not operate in this fashion. Behavior and brains influence each other; brains are dynamic and constantly in flux; and behavior is the outcome of a range of responses to stimuli. Brain activity is a global phenomenon, not merely a localized one within compartments of the brain, even for simple behaviors.

Most cognitive neuroscience studies of higher ordered behavior likewise assume a unidirectional view of behavior and brain activity: localized electrical impulses within the brain cause behavior, and therefore, bad behavior is always the consequence of brain activity beyond the conscious control of the person. Yet this view ignores a wealth of neuroscience evidence which shows that brains are dynamic and malleable with their structure and function readily changed by behavior itself.[22] To put it differently, behaviors influence brains inasmuch as brains determine behavior. This more integrated view of brains and behavior is of particular relevance to questions of culpability and responsibility where it is often assumed that a damaged brain should entail mitigation and excuse. On the contrary, if behavior influences brain structure and function, then under some circumstances, a damaged brain may evince willful engagement in deleterious behavior.

This more nuanced view of the brain stands as a hurdle for those who eagerly suggest dispensing with established

21. *See* WILLIAM R. UTTAL, NEUROSCIENCE IN THE COURTROOM 15–26 (2009).

22. As it was once assumed brains did not change structurally or functionally in adulthood, it is now understood that brains are constantly in flux through a process known as neuroplasticity. *See* Schwartz et al., *supra* note 4 *passim*; s*ee generally* Alvaro Pascual-Leone et al., *The Plastic Human Brain Cortex*, 28 ANN. REV. NEUROSCIENCE 377, 378 (2005) ("[P]lasticity is an intrinsic property of the nervous system retained throughout a lifespan and . . . The brain, as the source of human behavior, is by design molded by environmental changes and pressures, physiologic modifications, and experiences.").

doctrines of culpability and responsibility. In their zeal to adopt a culpability and punishment regime that relinquishes desert in favor of incapacitation and prediction of future dangerousness,[23] those who embrace the neuro-person model seem all too willing to accept a view of personhood that invariably trivializes the ability of individuals to exert control over their own behaviors in favor of one which reduces humanity to the indiscriminate ebb and flow of chemicals between neurons.[24] The fact that even those chemical messengers remain poorly understood appears to do little to curb the enthusiasm for this breathtaking endeavor.[25]

And that enthusiasm seems odd given the implications of the neuro-person model. While supporters grudgingly admit that biological science has an ugly history when applied to the law, few harbor fears that such mistakes would repeat themselves.[26] Instead, most neurolaw talk embraces the promise of changing minds by changing brains through a therapeutic model of criminal justice that views retributive punishment as inhumane, but classifying and modifying brains by their state as a great triumph of rehabilitation. As one leading scholar suggested, there is no real difference between prison and lobotomy.[27]

---

23. *See, e.g.*, Slobogin, *supra* note 9, at 122 ("The criminal law ought to embrace the dangerousness criterion, with the significant caveat that it do so wholeheartedly. . . .").

24. *See generally* Jones & Goldsmith, *supra* note 7, at 426 ("Synaptic connections in the brain ebb and flow–not only over the course of a lifetime, but during a single day."). While Jones and Goldsmith are undoubtedly correct in their assessment of synaptic physiology, this and similar biological facts do not sufficiently explain why people behave as they do. Rather, they merely describe physiological function. *Cf.* UTTAL, *supra* note 21, at 23–24 (detailing the fallacy that description equates an explanation).

25. *See, e.g.*, Theodore H. Bullock et al., *The Neuron Doctrine, Redux*, 310 SCI. 791, 792 (2005) (describing the shifting understanding of how neurotransmitters contribute to neuronal communications). Bullock and colleagues question "what features of the human brain account for our level of behavioral complexity? It is doubtful that the answer emerges from knowing the sheer number of cells, or the properties of synapses, or the identity of neurotransmitters and modulators." *Id.*

26. *See* Greely, *supra* note 8, at 1133 ("We should not view the fact that these possible interventions would intervene directly in a subject's brain as necessarily disqualifying them. Many . . . justifications for criminal sanctions work by affecting a criminal's brain."); Jones & Goldsmith, *supra* note 7, at 499 (acknowledging that science could be misapplied in a legal context, but arguing that societal fears should not bar the sensible use of scientific knowledge).

27. *See* Greely, *supra* note 8, at 1134 ("I see no qualitative difference

This credulous fervor for neurolaw hides its secret ambition. Those who view desert as an improvident distributive theory of punishment are many, but few suggest dispensing with the criminal justice system wholesale. Yet much of the neuro-person model's construction of culpability lies with the view that crime itself is a mental illness and not behavior of lawless citizens.[28] They consider crime the product of impaired brains and scientists are best suited for handling criminal justice policy, not lawyers. Once crime is understood as a behavioral problem rooted in the impaired brains of many unfortunate citizens, ameliorating crime will properly involve civil remedies instead of criminal ones. Therapeutic justice has already made substantial inroads under the reasonable view that criminal justice policy should encompass provisions which reduce offender recidivism. But in doing so, therapeutic justice seems all too eager in to jettison the adversarial system of justice in place of one which values intervention at the cost of adversarial rights and individual liberty by implementing adjudication forums that greatly diminish zealous advocacy by counsel. The very nature of the criminal justice system as one entrenched in laws which restrain government is quite foreign to therapeutic justice. And that is troublesome given the reemergence of the therapeutic state.

The balance of this Article is organized as follows. Part II explains the birth of neurolaw as a movement rooted in the construction of the mind by cognitive neuroscience and discusses the problems with that model of the mind. Part III examines the ambition and consequence of these constituent parts. The Conclusion offers some thoughts on neurolaw's implications.

II. THE BIRTH OF NEUROLAW

Neurolaw owes its fame to a modest and seemingly

between acting *directly* to change a criminal's brain–through drugs, surgery, DBS, or vaccines, if proven safe and effective–and acting indirectly–through punishment, rehabilitation, cognitive therapy, parole conditions–to achieve similar ends.").

28. *See generally* ADRIAN RAINE, THE PSYCHOPATHOLOGY OF CRIME: CRIMINAL BEHAVIOR AS A CLINICAL DISORDER (1997) (discussing the arguments and counterarguments for considering crime a disorder); Alec Buchanan & Howard Zonana, *Mental Disorder as the Cause of a Crime*, 32 INT'L. J. PSYCHIATRY & L. 142, 142–43 (2009) (arguing that psychiatry evidence should be used more in criminal proceedings).

innocuous beginning. Lawyer J. Sherrod Taylor coined the term during the early 1990s to describe the "converging courses" of neuropsychology and the legal system.[29] Adopted to explain the growing influence of expert testimony by neuropsychologists in brain-injury civil suits, neurolaw's initial focus rested on obtaining financial remedies for people with traumatic brain injuries.[30] As such, neurolaw had little ambition to modify criminal law doctrine or add to the debates of culpability and responsibility. Instead, neurolaw's domain lay chiefly with providing financial awards through civil litigation under the principle that monetary gains improved clinical outcomes for people with established brain injuries.[31]

But the allure of neurolaw from its conception was its ability to describe personhood by reference to structural and functional aspects of the brain. And this appeal was derived in large measure by powerful new technologies which gave unprecedented access to living brains. Brain imaging proved immensely useful not only for scientific inquiry but because the pictorial quality of the images it produced easily translated brains into entities that were accessible to non-scientists.[32] Even laypersons ponder at times how their brains manifest in the material world. Brain imaging allowed the non-neuroscientist to peer inside the brain's fantastic intricacy and discern a rudimentary understanding of how the human

29. *See* J. Sherrod Taylor et al., *Neuropsychologists and Neurolawyers*, 5 NEUROPSYCHOL. 293, 293 (1991).

30. *See* J. Sherrod Taylor, *Neurolaw: Towards a New Medical Jurisprudence*, 9 BRAIN INJ. 745, 746 (1995) ("The central thesis of neurolaw proposes that financial resources obtained through civil justice remedies contribute to improving the quality of life for persons with TBI [traumatic brain injury] and their families.").

31. *See* J. Sherrod Taylor, *An Overview of Neurolaw for the Clinician: What Every Potential Witness Should Know*, 16 NEUROREHABILITATION 69, 69 (2001) ("[N]eurolaw provides that better legal outcomes promote better clinical outcomes for patients with neurological injury.").

32. Non-experts are more willing to find explanations accompanied by brain images as more creditable. *See* David P. McCabe & Alan D. Castel, *Seeing Is Believing: The Effect of Brain Images on Judgments of Scientific Reasoning*, 107 COGNITION 343, 349–50 (2008). Even without pictures of the brain, explanations prefaced by the words "neuroscience explains" persuade people to accept causal events. *See* Deena Skolnick Weisberg et al., *The Seductive Allure of Neuroscience Explanations*, 20 J. COGNITIVE NEUROSCIENCE 470, 475 (2008) (finding that even irrelevant neuroscience information leads non-experts to prefer and accept explanations of behavior even when there are salient problems with the explanations as a whole).

experience involves an organ weighing about 3 pounds.[33]

The ability of this new neuroscience to make the complex simple is also what made it ideal for the law. Science and law are often said to exist on different dimensions: law operates chiefly through deduction, logic, and precedent; science is based on the scientific method, which values empirical and measurable phenomena as its centerpiece.[34] As such, science is increasingly multifaceted as the wealth of scientific information is exponential: we know much more about the world now than we did in the past. The consequence of this overflow of information means that the nuances of scientific explanations when applied to legal questions are often in tension with the established precedents of legal doctrine.

The simplicity of brain scanning images promises to overcome that difficulty. Ageless questions of how and why people think and behave as they do are considered fully explainable under the purview of neuroimaging by its supporters. The access that neuroimaging provides to living brains and the assumption that brain states equates all mentation easily wed it to a variety of legal questions.[35] Within the past several years, that road has ventured to the foundational tenets of responsibility and culpability. As neurolaw questions the very idea that people can choose their behavior, it inevitably questions the legitimacy and role of punishment.

A. BRAINS WITHOUT MINDS

The draw of neurolaw is a gloss of intrigue and seduction. Why people behave the way that they do is a perpetual and universal question. That question has often been understood to

33. *See* PAUL GLEES, THE HUMAN BRAIN 102 (2005).

34. *See* Holloway v. United States, 148 F.2d 665, 667 (D.C. Cir. 1945) ("The modern science of psychology is concerned with diagnosis and therapeutics and not with moral judgments. It proceeds on an entirely different set of assumptions [than the criminal law does]."); Gino C. Speranza, *The Medico-Legal Conflict over Mental Responsibility*, 13 GREEN BAG 123, 124–25 (1901) ("[L]aw and medicine represent distinct currents or forces of thought . . . .").

35. Most scholarship has centered on the use of neuroimaging in deception detection, yet there are numerous methodological problems with even this straightforward use. *See*, Sean A. Spence, *Playing Devil's Advocate: The Case* Against *fMRI Lie Detection*, 13 LEGAL & CRIMINOLOGICAL PSYCHOL. 11, 22–24 (2008).

entail some subjectivity across individuals and cultures.[36] Beyond the basilar needs of appetites and social bonding, diverging views of what causes people to behave as they do are copious and disputed.[37] Even within the behavioral sciences, there is much disagreement about why people behave as they do. For many years, the Freudians persuaded scientists and non-experts alike that behavior was the product of psychosexual drives established mostly during early childhood under the command of an immaterial unconsciousness.[38] In succeeding years, behaviorism proposed that external behaviors *were the mind* and that the mechanics operating within the brain were irrelevant.[39] Other theories abound but most hold to the immutable premise that people direct their behavior at least some of the time. Even under the subterfuge of the unconscious, psychoanalysis claimed that people acted because they had reasons—even when they were unaware of them.

The path of cognitive neuroscience stands in contrast. The foundational walls upon which it rests hold unwaveringly to the tenets of classical physics, reductive materialism, and hard determinism. Implicit in this model is the notion that, in time, all human experiences will be accessible by various physical apparatuses designed to explore the brain, that all mentation will be measurable by these devices, and that accurate predictions of future behavior by way of brain activity can be made solely by understanding the material properties of the brain.[40] While alternative psychological theories hold to some of these premises, cognitive neuroscience confidently suggests our perception of personhood grounded in the sense that we choose how to act is false and untenable.[41] Instead we are automatons,

---

36. *See* Paulo Sousa, *On Folk Conceptions of Mind, Agency and Morality*, 6 J. COGNITION & CULTURE 1 (2006) (discussing the connection between folk conceptions of mind, agency, and morality and scientific scrutiny of those conceptions).

37. *Id.* at 9–14 (discussing different scholars' interpretations of the folk concept of free will).

38. *See* STEPHEN A. MITCHELL & MARGARET J. BLACK, FREUD AND BEYOND 13–22 (1995).

39. *See* WILLIAM M. BAUM, UNDERSTANDING BEHAVIORISM31–40 (1994).

40. *See* Greene & Cohen, *supra* note 3, at 1781.

41. *Id.* at 1775 ("Cognitive neuroscience, by identifying the specific mechanisms responsible for behaviour, will vividly illustrate what until now could only be appreciated through esoteric theorizing: that there is something fishy about our ordinary conceptions of human action and responsibility.").

fooled by a belief in goal-directed behavior that we perceive is under our control but is entirely the product of forces set into motion long before our existence.[42] That we may believe that we prefer and choose to indulge in chocolate ice-cream over vanilla is an illusion; instead, we are a passive audience to the electrical cadence of neuronal firings buried deep within our heads.[43]

But even that view would not be correct according to the adherents of the neuroscience model of humanity. As many prominent neuroscientists have claimed, there is no "you" as is commonly understood: The brain and the mind are synonymous.[44] As such, the very idea that we are passive observers of our determined mind is fallacious. What we perceive as the mind is nothing more than a cognitive adaptation established by our brains to allow higher-ordered behavior.[45] That our brains engage in behavior before we become consciously aware of it means behavior operates independently from our consciousness.[46] At the least we are

42. *See id.* at 1777. *But see* LEWIS F. PETRINOVICH, HUMAN EVOLUTION, REPRODUCTION, AND MORALITY 112–15 (1995) (discussing problems with genetic determinism when applied to evolutionary psychology); Schwartz et al., *supra* note 4, at 1313–15 (2005) (arguing that a strict determinist view of behavior is incompatible with quantum physics).

43. *See* Martha J. Farah & Andrea S. Heberlein, *Personhood and Neuroscience: Naturalizing or Nihilating?*, AM. J. BIOETHICS, Jan. 2007, at 37, 40 ("The real contribution of neuroscience to understanding personhood may be in revealing not what persons are, but rather why we have the intuition that there are persons. . . . [P]ersonhood is illusory, constructed by our brains and projected onto the world."); *see generally* DANIEL M. WEGNER, THE ILLUSION OF CONSCIOUS WILL 2–28 (2003) (debating the existence of free will versus determinism).

44. *See* CRICK, *supra* note 5, at 3; Greene & Cohen, *supra* note 3, at 1778–9.

45. *See* COLIN BLAKEMORE, THE MIND MACHINE 270 (1988) ("We feel ourselves, usually, to be in control of our actions, but that feeling is itself a product of the brain. . . .").

46. This refers to the famous experiment that purported to show that the motor cortex is activated several hundreds of milliseconds before subjects become aware of their desire to act. Benjamin Libet et al., *Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential)*, 106 BRAIN 623, 635–36 (1983). *But see* HENRIK WALTER, NEUROPHILOSOPHY OF FREE WILL 250–52 (Cynthia Klohr trans., 2001) (discussing criticisms of Libet's interpretation of the experimental findings); Jing Zhu, *Reclaiming Volition: An Alternative Interpretation of Libet's Experiment*, 10 J. CONSCIOUSNESS STUD. 61, 67–74 (2003) (providing an alternative interpretation of Libet's results based on artifacts in Libet's experimental design and on the results of experiments by other investigators).

fools under the direction of our selfish genes;[47] at the worst our identity is utterly an illusion.[48] The very notion of human agency—that people evaluate their environments, make choices, and impose those choices in the world—is entirely incompatible with the cognitive neuroscience theory of personhood.[49] And that view has serious implications for theories of culpability and responsibility so fundamentally rooted in most legal systems.

The very idea that people are mere passive observers of the world in which they occupy is no recent development.[50] And while most people scoff at the idea that they have little, if any, control over their own behavior, Western culture has subtly trended towards that idea for many decades. Our modern culture is replete with notions of everyday behavior as products of addictions and compulsions. The very idea that how people think and evaluate the world is influenced—often quite strongly—by unconscious biases or motivations is a direct descendent of modernity's infatuation with behavioral science explanations of humanity. Irrespective of the veracity of this view, it is undeniable that modern culture looks first to psychologists and psychiatrists when seeking explanations for why people behave badly.[51]

Indeed, mental health professionals have made tremendous inroads in how most people think about the mind and behavior. In many ways, this is beneficial insofar as modern society has accommodated the view that people can change their unwanted behavior or relinquish their mental suffering by seeking professional mental health treatment. Likewise, there is a greater understanding that our brains do matter in terms of how our minds operate. Those who bear substantial brain defects are considered unlucky inhabitants of a broken physical body and not the wellspring of malevolent

---

47. *See* RICHARD DAWKINS, THE SELFISH GENE *passim* (1976).

48. *See* Farah & Heberlein, *supra* note 43, at 45 ("[P]ersonhood is a kind of illusion.").

49. *See* Steven P.R. Rose, *Human Agency in the Neurocentric Age*, 6 EMBO REP. 1001, 1001 (2005).

50. For an overview of theory of determinism see Peter Van Inwagen, *The Incompatibility of Free Will and Determinism*, 27 PHIL. STUD. 185, 185–88 (1975).

51. Perhaps the best example of this phenomenon is the numerous expert commentaries that are sought by the popular press when a heinous crime is committed. *See, e.g.*, Susan Hansen, *The Mind of a Killer*, SALON, Jul. 27, 2001, http://www.salon.com/books/int/2001/07/27/killers/.

forces or descendents of cruel or indifferent parents.[52] Modern psychotherapy and biological psychiatry have undoubtedly improved the lives of many simply because they offer hope through treatment to those whose prognosis decades ago would have been very poor.

But there is a dark side to this fascination with behavioral science. As there remains no definitive laboratory test for most mental disorders, their presence is assessed solely by descriptions written and endorsed by mental health professionals. That is, illnesses such as depression or anxiety are not diagnosed by a laboratory test of brain pathology, rather they are declared so based upon diagnostic descriptions of behaviors and reported emotional states.[53] Thus, someone is diagnosed with clinical depression when she reports feeling sad most days, cannot concentrate, and appears to engage in tasks with undue torpidity. This process is hardly unusual in medicine as many diagnoses begin with a clinical interview of reported symptoms by the patient. Yet behavioral science is vastly different because in most instances that is where the process ends. Despite the promise of bio-markers, genetic testing, and brain imaging, mental health practice is nearly devoid of these purposed confirmatory tools.[54] And while many are confident that technology will eventually overcome this limitation,[55] technology is unlikely to tell us what constitutes a normal mind.

The perennial question of what defines a normal mind scarcely constricts the continued encroachment of diagnostic labels for an endless number of behaviors and mental states.

---

52. Many influential authors during the zenith of psychodynamic thought concluded that parental behavior resulted in mental illness in children—including psychosis. *See, e.g.*, THEODORE LIDZ ET AL., SCHIZOPHRENIA AND THE FAMILY 15–19, 428–30 (1965).

53. A somewhat immoderate version of this thesis was famously introduced many decades ago in the wake of revelations that Russian psychiatrists often abused the powers of psychiatry to quell political dissidents. *See* THOMAS S. SZASZ, THE MYTH OF MENTAL ILLNESS 1–15, 62–72 (1961).

54. *See* Ed Bullmore et al., *Why Psychiatry Can't Afford to be Neurophobic*, 194 BRIT. J. PSYCHIATRY 293, 295 (2009) (stating that there is currently "no clear role for neuroimaging, biomarkers or genetic testing" in clinical psychiatry).

55. *See* Peter Tyrer, *Civil War Psychiatry*, 194 BRIT. J. PSYCHIATRY 386, 386 (2009) (claiming "in 20 years time it may be very different" in terms of the clinical application of neuroimaging, biomarkers, and genetic testing).

Over roughly the past fifty years, the number of official diagnosable mental disorders has increased by almost 300%.[56] That number will surely grow with the forthcoming fifth edition of the official diagnostic manual[57]—aided by the contributions of neuroscience.[58] Much of this growth can be attributed to the classification of behaviors once considered indicative of poor character or the willful immoderation of illicit passions. Intemperance, pride, and gluttony are no longer personal flaws deserving of shame but symptoms of interment-explosive disorder,[59] Narcissistic Personality Disorder,[60] and the various eating disorders requiring professional treatment.[61] An inappropriate concern with eating healthy foods is now indicative of a metal disorder.[62] So too, are a host of behaviors centered on poor self-control: overuse of the internet,[63] excessive sun-tanning,[64] and indulgent cell phone use[65] are

56. *See* Steven K. Erickson, *The Myth of Mental Disorder: Transsubstantive Behavior and Taxometric Psychiatry*, 41 AKRON L. REV. 67, 113–14 (2008)

57. *See* Roger K. Blashfield & Kenneth A. Fuller, *Predicting the DSM-V*, 184 J. NERVOUS & MENTAL DISEASE 4 (1996) (predicting 1,800 unique diagnostic criteria for the new edition based on growth from prior editions).

58. *See generally* Steven E. Hyman, *Can Neuroscience Be Integrated into the DSM-V?*, 8 NATURE REVIEWS NEUROSCIENCE 725, 728–31 (2007) (arguing that neuroscience will and should influence the forthcoming psychiatric diagnostic manual).

59. *See* AM. PSYCHIATRIC ASS'N, DIAGNOSTIC AND STATISTICAL MANUAL FOR MENTAL DISORDERS FOURTH EDITION: TEXT REVISION 663–67 (4th ed. 2000) [hereinafter DSM-IV REVISION].

60. *Id.* at 714–17.

61. *Id.* at 583–95.

62. *See* L. M. Donini et al., *Orthorexia Nervosa: A Preliminary Study With a Proposal for Diagnosis and an Attempt to Measure the Dimension of the Phenomenon*, 9 EATING & WEIGHT DISORDERS 151, 151 (2004).

63. *See* Jerald J. Block, *Issues for DSM-V: Internet Addiction*, 165 AM. J. PSYCHIATRY 306 (2008) (arguing that internet addiction is associated with tolerance and withdrawal symptoms and that it should be included in DSM-V). *But see* Nicki A. Dowling & Kelly L. Quirk, *Screening for Internet Dependence: Do the Proposed Diagnostic Criteria Differentiate Normal from Dependent Internet Use?*, 12 CYBERPSYCHOL. & BEHAV. 21, 22 (2009) (suggesting that current methods for diagnosing internet addiction do not discriminate normal from abnormal users).

64. *See* Molly M. Warthan et al., *UV Light Tanning as a Type of Substance-Related Disorder*, 141 ARCHIVES DERMATOLOGY 963, 965 (2005) (arguing that excessive sunbathing is an addiction and should be regulated by the law).

65. *See* Lauren D. LaPorta, *Cellular Telephones: A New Addiction?*, PSYCHIATRIC TIMES, Oct. 1, 2006, http://www.psychiatrictimes.com/display/article/10168/52076?pageNumber=1.

considered forms of addictions necessitating inventions by mental health professionals. Very few people escape the increasing reach of psychiatric diagnostic classification. In a few short decades, we all have become mentally unsound.

B. THE WAY OF COGNITIVE NEUROSCIENCE

Cognitive neuroscience can be defined as an investigative field that seeks to understand how the mind arises from the central nervous system, in particular the brain.[66] It bridges the fields of cognitive science, psychology, and biology by focusing on the biological mechanisms underlying cognition, with a specific focus on the neural substrates of mental processes and their behavioral manifestations.[67] It claims to answer how psychological and cognitive capacities are produced by the structure and function of neural networks within primates.[68] It is premised on the idea that all aspects of mentation are entirely reducible to material terms and these elements are measurable in ways that provide meaningful insights into the origin of behaviors.[69] It assumes that measured brain states can be used not only to understand how brain activity becomes behavior, but that such information can be generalized across all brains to arrive at a unified understanding of human mentation.[70] To put it differently, cognitive neuroscience asserts that all individual thoughts, emotions, and feelings can be traced to certain defined biological locations of the brain.[71] Moreover, these measurements can be used to discern how all brains operate and how to build models which predict future mental activity and behavior.[72]

It is hard to imagine the domain of cognitive neuroscience without reference to the various brain imaging technologies

66. *See* HOWARD S. KIRSHNER, BEHAVIORAL NEUROLOGY 5 (2d ed. 2002).

67. *See id.* at 4–5.

68. *See generally* David Wells, Book Review, 122 BRAIN 2413 *passim* (1999) (reviewing RANDOLPH W. PARKS ET AL., FUNDAMENTALS OF NEURAL NETWORK MODELING: NEUROPSYCHOLOGY AND COGNITIVE NEUROSCIENCE (1999)) (overviewing a textbook that explains neural network modeling).

69. *See* Martha J. Farah, *Neuroethics: The Practical and the Philosophical*, 9 TRENDS IN COGNITIVE SCI. 34, 34 (2005).

70. *Id.*

71. *Id.* at 38.

72. *Id.*

that it so frequently utilizes to arrive at its conclusions.[73] These technologies have greatly enhanced the ability of scientists to investigate the structure and function of the living brain. The principal technology that cognitive neuroscience uses is functional magnetic resonance imaging (fMRI), which began to dominate the field of brain mapping during the 1990s.[74] Like its predecessor, structural magnetic resonance imaging (MRI), fMRI uses a powerful magnetic field known as Larmor precession to affect the water molecules within cells.[75] While the technical details of brain imaging are beyond the scope of this article, the key difference between structural and functional MRI lies in what they purport to measure and explain. As its name implies, structural MRI provides images of brain structure, that is, how the brain is organized by its constituent parts.[76] The pictures provided by structural MRI conform to what is understood about how the brain is constructed and organized when it is dissected at autopsy.[77] The chief advantage of structural MRI is, of course, that it provides an examination of structure inside *living* brains. In contrast, functional MRI claims to show how brains function and work.[78] As such, the latter imposes two significant assumptions absent from the former. First, since there are no definitive or universal rules about how the brain becomes the mind,[79] fMRI utilizes technologically convenient assumptions about brain function to arrive at its end product.[80] Second, there are no alternative mechanisms equivalent to autopsy

73. For example, a recent review discovered approximately 19,000 peer-reviewed articles involving fMRI technology since 1991, corresponding to over three papers per day. Nikos K. Logothetis, *What We Can Do and What We Cannot Do with fMRI*, 453 NATURE 869, 869 (2008).

74. *See id.*

75. RAY H. HASHEMI, ET AL., MRI: THE BASICS 26 (2003).

76. *See* DONALD W. MCROBBIE ET AL., MRI: FROM PICTURE TO PROTON 1 (2003).

77. *See id.*

78. MCROBBIE ET AL., *supra* note 76, at 333.

79. *See* Stephen J. Morse, *Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience*, 9 MINN. J. L. SCI. & TECH. 1, 20 (2008) ("We do not know how the brain enables the mind . . . ."); UTTAL, *supra* note 21, at 27 ("The mind, however difficult it may be to define, is widely agreed to be the outcome of complex information processing associated with the brain . . . . [M]ost scientists and philosophers also agree that how this is accomplished remains mysterious and unknown.").

80. WILLIAM R. UTTAL, THE NEW PHRENOLOGY: THE LIMITS OF LOCALIZING COGNITIVE PROCESSES IN THE BRAIN 136 (2001) [hereinafter UTTAL, LOCALIZING].

dissection to validate the findings of fMRI. Hence, it remains difficult, if not impossible, to replicate fMRI findings outside of the purview of the brain imaging machine. Thus, there is a leap of faith built into the brain imaging framework when we accept its conclusions regarding the operations of the mind. Since cognitive neuroscience is ultimately interested in how the mind operates through the brain,[81] the assumptions built into the fMRI model matter immensely as fMRI is the workhorse of the field.

A fundamental attribute of the fMRI technique holds that understanding the complexity of mentation is possible by reducing the operation of the brain to multiple quasi-independent cognitive modules.[82] For instance, sight is assigned to regions of the brain believed responsible for generating and regulating vision, bodily movement to areas involved with motor control, and so forth.[83] Consequently, when researchers want to explore how people see something, they examine those parts of the brain which are believed involved with sight and usually exclude other regions of the brain. The images produced by the fMRI technique rest upon the notion that subtracting fMRI data at different time points from areas of the brain presumed to be involved with particular elements of thinking and behavior reveals meaningful differences between brain states that conveys valuable information about human mentation.[84] That is, functional brain imaging concerns

81. *See* Greene & Cohen, *supra* note 3, at 1781 ("There are many causes that impinge on behaviour, but all of them—from the genes you inherited, to the pain in your lower back, to the advice your grandmother gave you when you were six—must exert their influence though the brain.").

82. Known as "localization," this theory is the backbone of cognitive neuroscience and holds that mental functions are carried out by particular brain structures. *See* EDWIN CLARKE ET AL., *The Genesis of Cortical Localization*, *in* AN ILLUSTRATED HISTORY OF BRAIN FUNCTION 115, 115–28 (1996); *cf.* WILLIAM R. UTTAL, NEURAL THEORIES OF MIND 116 (2005) [hereinafter UTTAL THEORIES] ("On the psychological side, it is not at all certain that cognitive modules exist."); Logothetis, *supra* note 73, at 869 ("[A] frequently made assumption is that the mind can be subdivided into modules or parts whose activity can then be studied with fMRI. If this assumption is false, then even if the brain's architecture is modular, we would never be able to map mind modules onto brain structures . . . .").

83. JOSEPH E. LEDOUX, THE EMOTIONAL BRAIN: THE MYSTERIOUS UNDERPINNINGS OF EMOTIONAL LIFE 76–77 (1996).

84. Brain imaging uses the technique of generating differences between groups by subtracting brain activation patterns to arrive at its conclusion of individual differences. For an overview, *see* Matthew B. Crawford, *The Limits*

itself with the idea that areas of the brain that consume more oxygenated blood than others do while performing a task discloses how people think.[85] This theory of brain activity is harmonious with the structural conception of the brain whereby different areas of the brain control different activities. Thus, the frontal lobe of the brain is considered involved with conscious thought,[86] the temporal lobe with smell[87] and sound, and the pituitary with hormone regulation.[88] Modern neuroscience arrived at this global understanding of brain function mainly by observing functional and behavioral deficits among those with brain injury.[89] Those who, for instance, suffer from traumatic damage to the occipital lobe frequently endure visual impairments, including hallucinations or even blindness.[90]

When neuroscience explores functions such as sight, smell, or motor coordination, it operates under the reasonable assumption that outward behavior is the product of brain activity controlled by specific regions of the brain. But the chief interest of cognitive neuroscience lies beyond observable behavior. Instead of explaining behavior, cognitive neuroscience seeks to understand how people *think*. To accomplish this feat, it must employ experiments which assume

---

of Neuro-Talk, 19 NEW ATLANTIS, Winter 2008, at 65, 70-72. These group differences often obscure tremendous variations among the individual subjects. *See* Michael B. Miller et al., *Extensive Individual Differences in Brain Activations Associated with Episodic Retrieval Are Reliable over Time*, 14 J. COGNITIVE NEUROSCIENCE 1200, 1209–11 (2002). Additionally, the calculations used with this technique have recently been criticized for generating inflated estimates of significance. *See* Edward Vul et al., *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition*, 4 PERSP. ON PSYCHOL. SCI. 274 *passim* (2009); Nikolaus Kriegeskorte et al., *Circular Analysis in Systems Neuroscience: The Dangers of Double Dipping,* 12 NATURE NEUROSCIENCE 535 *passim* (2009).

85. It is worth noting that only recently was it discovered that cells in the brain were responsible for the increased blood flow measured by fMRI. *See* James Schummers et al., *Turned Responses of Astrocytes and Their Influence on Hemodynamic Signals in the Visual Cortex*, 320 SCI 1638 (2008).

86. KIRSHNER, *supra* note 66, at 22–23.

87. EVIAN GORDON, INTEGRATIVE NEUROSCIENCE: BRINGING TOGETER BIOLOGICAL, PSYCHOLOGICAL AND CLINICAL MODELS OF THE HUMAN BRAIN 182 (2000).

88. Marianne B. Müller et al., *Genetics of Endocrine—Behavior Interactions, in* 5 HORMONES, BRAIN, AND BEHAVIOR 263 (Donald W. Pfaff ed., 2002).

89. *See* GORDON, *supra* note 87.

90. *See* ORRIN DEVINSKY & MARK D'ESPOSITO, NEUROLOGY OF COGNITIVE AND BEHAVIORAL DISORDERS 132–38 (2004).

that complex human thoughts are reducible to independent modules located in specific areas of the brain.[91] Thus, cognitive neuroscience claims that when one is faced with an important moral dilemma, such as whether to pull a lever to divert a trolley car and save many lives at the expense of killing another,[92] how one thinks about such morally impinged matters can be *entirely* understood by examining brain activity in certain sections of the brain.[93] Yet this approach assumes that such complex thoughts are accessible and readily explained by the technology employed by the neuroscientist. Determining which areas of the brain are responsible for such higher-ordered, and presumably uniquely human mentation, is largely a matter of reliance on three factors: (1) anecdotal evidence from brain injury victims, (2) observations of people under the effects of psychoactive chemicals, (3) and highly-contrived experiments conducted within the confines of fMRI laboratories.

1. Reduction and Deduction

Anecdotal evidence of brain function related to higher cognitive processes comes from retrospective studies that link severe and substantial brain damage to observed behavior.[94] Perhaps the most famous example of this phenomenon was the nineteenth century case of Phineas Gage, a law-abiding railway

91. *See* UTTAL, LOCALIZING, *supra* note 80, at 99–110; *see also* Daniel T. Willingham & Elizabeth W. Dunn, *What Neuroimaging and Brain Localization Can Do, Cannot Do, and Should Not Do for Social Psychology*, 85 J. PERSONALITY & SOC. PSYCHOL. 662, 667–68 (2003) (describing the difficulty of localizing social psychology constructs).

92. This, of course, refers to the legendary trolley car dilemma. *See* PHILIPPA FOOT, VIRTUES AND VICES AND OTHER ESSAYS IN MORAL PHIL. 19, 20–24 (1978); *see also* Judith Jarvis Thomson, *The Trolley Problem*, 94 YALE L. J. 1395 *passim* (1985) (providing an extended discussion of the dilemma); Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCI. 2105 *passim* (2001) (investigating emotional processing areas of the brain using the trolley car dilemma and fMRI).

93. *Cf.*Steven Pinker, *The Moral Instinct*, N.Y. TIMES MAG., Jan. 13, 2008, at 32–34 ("The human moral sense turns out to be an organ of considerable complexity, with quirks that reflect its evolutionary history and its neurobiological foundations.").

94. S*ee, e.g.*, Thomas C. Neylan, *Frontal Lobe Function: Mr. Phineas Gage's Famous Injury*, 11 J. NEUROPSYCHIATRY & CLINICAL NEUROSCIENCE 280, 280 (1999) (explaining how a physician used Phineas Gage's injury and reported personality changes to advance the idea of cerebral localization in the 1800s).

worker who suffered a severe brain injury when an iron tapping rod was driven though a large portion of his frontal lobes.[95] After the accident, Gage's behavior was erratic and callous; friends who knew him noted his changed personality despite his largely preserved cognitive abilities.[96] Gage's injury in many ways was a watershed moment in medicine. It prodded a skeptical medical community to conclude that personality, behavior, and the brain were inextricably linked.[97] It also hastened the view that intrinsic elements of personality and behavior were localized within discrete areas of the brain and laid open the possibility that surgical intervention could radically alter sensation, perception, behavior—and presumably thoughts as well.[98] The hunt for connections between brain regions and defined behavior was launched; and that pursuit continues with vigor to this day.[99]

Much like the discoveries surrounding Phineas Gage, serendipity played a large role in revealing how psychoactive chemicals affect behavior. From the late nineteenth century through the 1930s, a number of French and German scientists were searching for new types of dyes to stain slides for microscopes.[100] These phenothiazine dyes were noted for their various medicinal properties, including the treatment of malaria during World War I, when the traditional quinine treatment became unavailable due to military blockades.[101] In the 1930s, phenothiazines were explored for their possible antihistaminic properties in an effort to discover new treatments for surgical shock.[102] After several formulations, the

95. James M. Harlow, Letter to the Editor, *Passage of an Iron Bar Through the Head*, 39 BOSTON MED. & SURGICAL J. 389–93 (1848), *reprinted in* 11 J. NEUROPSYCHIATRY & CLINICAL NEUROSCIENCE 281, 281–83 (1999).

96. *See* Neylan, *supra* note 94.

97. *See* MALCOLM MACMILLAN, AN ODD KIND OF FAME 41–45 (2002).

98. *Id.* at 125–42, 205–50. As might be expected, the lobotomy procedure was a direct descendent of this insight into brains, behavior and surgical intercession. *See* Robert P. Feldman & James T. Goodrich, *Psychosurgery: A Historical Overview*, 48 NEUROSURGERY 647, 649–54 (2001).

99. *See* Logothetis, *supra* note 73, at 869, 870 (claiming that about half of the 19,000 peer-reviewed papers utilizing fMRI technology explored "functional localization," of which many used lower powered fMRI magnets prone to "localization errors").

100. *See* Francisco López-Muñoz et al., *History of the Discovery and Clinical Introduction of Chlorpromazine*, 17 ANNALS OF CLINICAL PSYCHIATRY 113, 114 (2005).

101. *Id.*

102. *Id.* at 116.

drug chlorpromazine was discovered and used in surgery by a French army surgeon named Henri-Marie Lamborit.[103] Lamborit noted that chlorpromazine was a highly effective sedative and soon suggested to his psychiatric colleagues that they try it with psychiatric patients.[104] Its effect was monumental. Chlorpromazine would become the first effective psychiatric medicine and was largely responsible for the massive decline of institutional state asylums during the 1960s.[105] Exploring the mechanisms behind chlorpromazine and similar drugs led to understanding that neurotransmitters—the chemical messengers of neural tissue—can be affected by drugs to produce profound behavioral changes in people.[106] Those changes can both produce beneficial and horrific behaviors and suggest that behavior is directly linked with the chemistry of the brain. Those effects also imply that behavior results from the actions of neural mechanisms within defined regions of the brain.

The spate of brain imaging studies that claim to identify regions of the brain responsible for character, love, morality, and other human attributes also rely upon the premise that localized areas of the brain are responsible for human behaviors.[107] Many of these studies depend on comparisons between people with known brain injuries and those with no record or indication of brain injury. For instance, numerous studies have suggested that one area—the ventromedial prefrontal cortex (VMPC)—is predominantly responsible for moral behavior in humans.[108] These studies draw their conclusions by having subjects engage in simplified tasks or merely having them observe visual stimuli within the housing of the fMRI machine. For instance, one typical study asked subjects to choose between two given alternatives in a series of

103. *Id.* at 116–26.

104. *Id.* at 118.

105. *See id.* at 128.

106. *See* Solomon H. Snyder et al., *Drugs, Neurotransmitters, and Schizophrenia*, 184 SCI. 1243 (1974).

107. *See* Martha J. Farah & Nancy Murphy, *Neuroscience and the Soul*, 323 SCI. 1168, 1168 (2009) ("[A]s neuroscience begins to reveal the mechanisms underlying personality, love, morality, and spiritualty, the idea of a ghost in the machine becomes strained. Brain imaging indicates that all of these traits have physical correlates in brain function.").

108. *See, e.g.,* Michael Koenigs et al., *Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments*, 446 NATURE 908, 908–10 (2007).

hypothetical moral dilemmas.[109] The type of dilemmas used in these studies range from choosing environmental polices based on mortality calculations to cheating on one's taxes, to smothering a crying baby to prevent discovery of a group of people hiding from enemy soldiers.[110] Another study scanned subjects' brains while they observed emotionally laden facial features projected on to a television screen inside the fMRI apparatus.[111] Still another study examined probabilistic judgments in children assessed as having psychopathic traits while scanning their brains.[112] While the methods of these studies differ, all of the studies claim to link impaired moral judgment and emotional regulation based on differences in purported measured brain activity. Moreover, they all propose that the results generated by these tightly controlled studies, which take place in the unnatural environment of the fMRI machine, elucidate more generally how people behave in the world under the demands of everyday life.[113] Despite the small sample sizes used in these studies, primarily ranging from ten to twenty subjects,[114] supporters ardently suggest that these differences in brain activity demonstrate the root origin of behavior and thought.[115] Thus, moral thinking and behavior is

---

109. *Id.* at 108.

110. *See* Joshua D. Greene et al., *Supplement Data: The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 14 NEURON S1, S3–4 (2004), http://download.cell.com/neuron/mmcs/journals/0896-6273/PIIS0896627304006348.mmc1.pdf

111. *See* Andrea S. Heberlein et al., *Ventromedial Frontal Lobe Plays a Critical Role in Facial Emotion Recognition*, 20 J. COGNITIVE NEUROSCI. 721, 723–24 (2008).

112. *See* Elizabeth C. Finger et al., *Abnormal Ventromedial Prefrontal Cortex Function in Children with Psychopathic Traits During Reversal Learning*, 65 ARCHIVES GEN. PSYCHIATRY 586, 587–89 (2008).

113. *See generally* Andrew Scull, *Minds, Brains, Law, and Culture*, 130 BRAIN 585, 589 (2007) (book review) ("[J]ust as economists traditionally rely upon absurdly oversimplified portraits of human motivation to construct their models, so all the neuroscientific findings that are so proudly proffered reflect simple simulated experiments that in no way capture the intricacies of everyday social situations . . . .").

114. A recent commentary suggests that these small sample sizes greatly inflate the statistical findings in these studies. *See* Tal Yarkoni, *Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power—Commentary on Vul et al. (2009)*, 4 PERSP. ON PSYCHOL. SCI. 294 *passim* (2009).

115. *But see* Greg Miller, *Growing Pains for fMRI*, 320 SCI. 1412, 1413 (2008) (reviewing criticisms of the moral dilemma fMRI studies and noting that "some of the 'emotional' brain regions in the morality study have also been connected to memory and language—a caveat that is rarely mentioned in

merely a product of neuronal discharges inside the ventromedial prefrontal cortex of the brain—and nothing more.[116] They are, as two preeminent neuroscientists put it recently, "not mere correlates but are the physical bases of these aspects of our personhood."[117]

2. The Life of the Brain

The evidence that neuroscience marshals in favor of its explanation of higher-ordered behavior is striking. Few would seriously contend that modern understandings of behavior have not advanced immensely by examining the principal proxy of behavior which is the brain. Yet the relationship between brains, behavior, and minds is highly intimate and masked.[118] It is not at all clear that mentation can be reduced in any meaningful fashion by merely describing the activation patterns of neuronal tissue.[119] While neuroscientists would

---

media coverage . . . .").

116. The proponents of this neurocentric view suggest these findings are fundamentally important when examining normative ethics as well. *See, e.g.*, Joshua Greene, *From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?*, 4 NATURE REVS. NEUROSCI. 847, 847 (2003) ("[S]cientific facts can have profound moral implications and . . . moral philosophers have paid too little attention to relevant work in the natural sciences.").

117. Farah & Murphy, *supra* note 107, at 1168.

118. As Nobel prize winning scientist Max Delbrück put it in a quote loosely attributed to him: "So far as I can tell, our science simply has no handle whatever on the most conspicuous and immediate reality of our lives: that we are aware." *See* MAX DELBRÜCK, MIND FROM MATTER? 159 (1986).

> "When I was a student in Gödttingen, there appeared a story in the newspaper about a rabbit that had been frozen and then brought back to life. Although the report of this cryogenic feat was probably a sensationalist fabrication, someone was inspired by this story to ask various people what they would ask about if they had been frozen for 500 years and then revived." My own priorities come to mind with crashing immediacy: "Has a machine yet been made that has awareness, a synthetic consciousness? Is there any agreed-upon way yet to detect and measure the quality, or even the existence, of an object's consciousness?"

Id.

119. *See* Scull, *supra* note 113, at 589.

> Much is made of the fact that particular regions of the brain show heightened levels of activity on fMRIs when people, for example, are making choices, or telling lies. . . . [S]uch correlations prove nothing about the causal process involved, any more than . . . the existence of a particular sequence of events demonstrates that some early event in the sequence ineluctably caused the a later event. Post hoc ergo propter hoc is an elementary logical fallacy.

have us believe that their brain activity measurements equate an adequate and sufficient detailing of the mind, the impressive images upon which their persuasion relies runs counter to the tremendous complexity of brain anatomy and function. "With 10^{10} neurons and 10^{14} connections in the [human] cortex alone," the human brain defies any simple mechanical explanation.[120] Much of the workings of the brain are breathtakingly beautiful, mysterious, and continue to confound researchers despite their dogged attempts to dissect and translate its processes. Even the technology used to generate functional brain images remains clouded in obscurity, as many of the fundamental technicalities remain poorly understood.[121]

More importantly, the underlying cognitive processes and mechanisms which cognitive neuroscience seeks to explain are neither directly observable nor measurable. That is, mentation is not measured by the neuroscientist; brain activity is.[122] As such, when the neuroscientist claims to measure "moral thinking," what is measured are regions of brain activation believed to be associated with thinking about moral problems. But in describing that process, neuroscience has not yet explained how molecular brain activity becomes thoughts about morality. It is not as though neuroscience has any instrument known as a "moral thought" aperture. Instead, it has sophisticated tools which measure the proxy of thoughts, which is brain activity. Putting aside the endless debate of dualism between mind and brain, what matters in terms of understanding cognitive neuroscience's many claims surrounding measured mentation is the simple fact that, as of yet, we have no idea how the brain *becomes* the mind.[123] Describing brain molecular activity does little to advance that understanding.[124] At the most, brain imaging tells us what

---

Id.

120. *See* Logothetis, *supra* note 73, at 871–72.

121. A number of articles have illuminated the limitations of fMRI technologies, ranging from measurement and the physiology of the brain to statistical analysis of the data to the underlying scientific assumptions used. *See e.g., id. passim*; Vul, *supra* note 84 *passim*.

122. *See, e.g.,* Schwartz et al., *supra* note 4, at 1310 ("[T]erms such as 'feeling,' 'knowing,' and 'effort,' because they are intrinsically mentalistic and experiential, cannot be described exclusively in terms of material structure.").

123. *See* Morse, *supra* note 79, at 20.

124. *Cf.* Rose, *supra* note 49, at 1004 ("Research over many decades has produced an account of the molecular cascade occurring during memory

chemical and molecular patterns emerge from the brain when engaged in a task. As such, it describes forms of brain activity; but those contours of chemical variation do not explain why people think pushing someone into the path of a trolley car is immoral. To say an area of the brain is activated in some people when faced with this morally fraught dilemma only exposes the fact that cognitive neuroscience defines its relevant explanations by the technology it employs.[125] And that technology operates under assumptions and limitations which circumscribe its explanations.

Likewise, cognitive neuroscientists conveniently ignore evidence that brains operate in a fashion opposed to the strictures of brain imaging technology. Brain activation and localization betray the essential truth that when brains are engaged in a task, the entire brain is activated.[126] The localized areas of activation which are prominently displayed in pages of most neuroscience journals are based on the premise that subtracting activation images between groups reveals distinct brain regions responsible for the thoughts and behaviors in question.[127] And while it is understood that certain brain regions are necessary and crucial for defined behaviors—such as the motor cortex for movement—those observations in no way tell us which areas are sufficiently responsible for thinking about running down to the store for a pint of milk, much less proscribed acts of predation. To assert that a brain function, such as visual attention, seems involved with the visual and parietal lobes of the brain[128] fails to elucidate how the mind is engaged when looking at the works of Cezanne, Rubens, or

---

formation. However, this summary does not explain the memory, it merely describes the brain events involved in making it.").

125. *See* Logothetis, *supra* note 73, at 869 ("[L]ike all haemodynamic-based modalities, [fMRI] measures a surrogate signal whose spatial specificity and temporal response are subject to both physical and biological constraints.").

126. *See* Crawford, *supra* note 84, at 70

> In the case of functional (as opposed to structural) neuroimaging, what you are seeing when you look at a brain scan is the result of a subtraction. . . . One . . . problem is that this method eliminates from the picture the more massive fact, which is that the entire brain is active in both conditions.

Id.

127. *See id.*

128. *See, e.g.*, Jody C. Culham et al., *Attention Response Functions: Characterizing Brain Areas Using fMRI Activation During Parametric Variations of Attentional Load*, 32 NEURON 737, 742 (2001).

even a blank sheet of paper. As psychologist William Uttal notes:

> Is it a "stuff" that can be divided, allocated, and focused and that is available only in limited amounts, and thus can be localized in a particular part of the brain? Or, to the contrary, is it an attribute or characteristic of perception . . . inseparable as the diameter or whiteness of a golf ball is from the physical ball itself? . . . It seems plausible that many of the psychological components or modules we seek to locate in a particular region of the brain should likewise be thought of as properties of a unified mental "object" rather than as analyzable and isolatable entities.[129]

What Uttal suggests is what emerging neuroscience evidence reveals to be so: that the brain is a highly distributed system with millions of connections at play in any given moment.[130] As such, how people think about matters of love, jealously, and morality entails understanding how thoughts, behavior, and the brain are highly interwoven. While brain imaging experiments—and particularly the interpretations derived from their results—often assert that thinking and human behavior are merely the end products of determined brain functions, other evidence suggests the human mind is much more than this automated view.

One of the key insights provided by neuroscience entails the dynamic nature of the brain. While it was once assumed that adult brain structure was static and mostly unaltered by the environment and behavior itself, it has been firmly established that this is not the case.[131] It is now known that the brain is constantly in flux, with its structure and function ceaselessly molded by environmental influences.[132] Among these environmental influences is thought itself, which has the propensity to alter a wide-range of brain structures and functions ranging from perception of sensory stimuli to higher-

---

129. UTTAL LOCALIZING, *supra* note 80; *cf.* Crawford, *supra* note 84, at 67–69 (dissenting from Uttal's description of attention as "stuff" by noting that most neuroscientists define attention as a function but "this correction does not vacate the force of Uttal's criticism, because functions, like properties, are distributed . . . .").

130. *See* Michael S. Gazzaniga, *The Law and Neuroscience*, 60 NEURON 412, 413 (2008) ("The brain is a highly parallel and distributed system with literally millions of decisions being made simultaneously.").

131. *Id.*

132. Peter Vestergaard-Poulsen et al., *Long-Term Meditation Is Associated with Increased Gray Matter Density in the Brain Stem*, 20 NEUROREPORT 170 (2009) (reporting changes in neuronal cell bodies among those who engage in sustained attention activities).

ordered behavior.[133] For instance, several studies have demonstrated that psychological therapies which center on having patients change their thinking about certain problematic behaviors produces significant changes in cerebral metabolism, neuronal structure, and regional brain function.[134] These studies suggest that thoughts can alter brain structure and function inasmuch as the reverse may do so. To put it differently, the mental phenomenon which is only indirectly measured in the brain imaging paradigm by examining proxies of neuronal activity may very well exert independent effects upon the neuronal activity itself. As it was recently put by a group of prominent neuroscientists:

> [T]he assumption that all aspects of mental activity and emotional life are ultimately explicable solely in terms of micro-local deterministic brain activity, with no superposed effects of mental effort, produces a theoretical structure that both fails to meet practical scientific needs, and also fails to accord with the causal structure of modern physics.[135]

Yet discussions of law and neuroscience are nearly devoid of this vital aspect of brain physiology. The general thesis employed by those who wish to advance neuroscience within criminal law suggest that abnormal brain function is indicative of an agent wholly at the mercy of his aberrant and determined brain activity.[136] These hapless citizens are said to be in no way responsible for their illicit behavior in the same way that a ball is not responsible for rolling down a hill.[137] The determined forces operating on the behavior of the ball and the person

---

133. *See* Antoine Lutz et al., *Long-Term Mediators Self-Induce High-Amplitude Gamma Synchrony During Mental Practice*, 46 PROC. NAT'L ACAD. SCI. 16371–72 (2004) (describing how mental training can produce short-term and long-term neural changes)*;* Vestergaard-Poulsen, et al., *supra* note 132.

134. *See* Vincent Paquette et al., *"Change the Mind and You Change the Brain": Effects of Cognitive-Behavioral Therapy on the Neural Correlates of Spider Phobia*, 18 NEUROIMAGE 401, 406–07 (2003) (reporting psychotherapy modified dysfunctional neuronal structure and function among those with diagnosed anxiety disorders).

135. Schwartz et al., *supra* note 4, at 1312.

136. Greene and Cohen make this familiar argument using their example of Mr. Puppet, a person biologically designed with a predilection towards violence and vice. They conclude "in a very real sense, we are all puppets. The combined effects of our genes and environment determine all of our actions . . . . We are no more free than he is." Greene & Cohen *supra* note 3 at, 1780.

137. *Id.* at 1779.

plunging a knife into his wife's chest are one in the same.[138] Both are helpless in avoiding their fates disposed by the material forces at work set in motion outside of their consciousness.[139] Indeed, it is not just those with broken brains who reside in this stark illusion of personhood, but all of humanity.[140]

III. NEUROLAW'S SECRET AMBITION

The law cares much about causation. Whether in the arena of elaborate regulation or intrinsic rules proscribing harm to others, scarcely any law exists without the idea that breach of a law involves action (or omission) by an agent. From this simple premise flows the notion of culpability. Western legal systems rest on the assumption that people, acting as agents of their own behavior (at the very least) are responsible for their decisions and behaviors.[141] The law justifiably holds those who violate rules as blameworthy and responsible since it is presumed within most legal codes that people are responsible for their actions.[142]

It can be fairly said that an elementary reason why people obey the law is because the system of rules embodied within it follows the ordinary intuitions of how behavior should be

---

138. *Id.*

139. *Id.* at 1777 ("[A] deterministic universe starts however it starts and then ticks along like clockwork from there . . . . [Y]our sense of yourself and others as having free will is an illusion.").

140. *Id.* at 1775 (suggesting that neuroscience will reveal "there is something fishy about our ordinary conceptions of human action and responsibility . . . .").

141. As it was famously put by Hale:

> Man is naturally endowed with these two great faculties, understanding and liberty of will, and therefore is subjected properly capable of a law properly so called, and consequently obnoxious to guilt and punishment for the violation of that law, which in respect of these two great faculties he has a capacity to obey. The consent of the will is that, which renders human actions either commendable or culpable, as there is no law there is no transgressions, so regularly there is no will to commit an offense, there can be no transgression or just reason to incur the penalty.

1 HALE, PLEAS OF THE CROWN 14–15 (1986); *see also* Smith v. Armontrout, 865 F.2d 1502, 1506 (1988) ("The whole presupposition of the criminal law is that most people, most of the time, have free will within broad limits.").

142. *See* HLA HART & A.M. HONORÉ, CAUSATION IN THE LAW 59 (1959) ("In the moral judgments of ordinary life, we have occasion to blame people because they have caused harm to others . . . in all legal systems, liability to be punished . . . depends on whether actions (or omissions) have caused harm.").

regulated in everyday life.[143] Those intuitions are firmly rooted within the social view that people evaluate their environments, make choices, and impose those choices to the best of their ability on the world.[144] The law has long recognized that those who are unable to make choices or operate under extreme duress at the time they make a choice are not legally culpable agents.[145] Thus, incapacitated people cannot commit to contracts as well as those who commit crimes when under the immediate threat of death from another are not held responsible for their behavior.[146]

In many ways, agency is the crucible of law and humanity. The power of law is derived from the near universal belief that people can be persuaded in some fashion to abide by its dictates or suffer punishment for failure to do so.[147] Irrespective of whether the law is inherently just or overly oppressive, law operates to influence behavior and the choices of those who live under its regime.[148] While scholars may debate the essence of justice and mercy, Western legal conceptions of these elements derive in large measure from the idea that people choose and intend their behavior most of the time.[149] To undo agency risks

143. *See* Paul H. Robinson, *Why Does the Criminal Law Care What the Layperson Thinks is Just? Coercive Versus Normative Crime Control*, 86 VA. L. REV. 1839, 1861 (2000). Even fervent neurolaw supporters admit as much. *See* Greene & Cohen, *supra* note 3, at 1778 ("The legitimacy of the law itself depends on its adequately reflecting the moral intuitions and commitments of society.").

144. *See* United States v. Lyons, 739 F.2d 994, 994 (1984) ("An adjudication of guilt is more than a factual determination that the defendant pulled a trigger, took a bicycle, or sold heroin. It is a moral judgment that the individual is blameworthy."); Alan R. Felthous, *The Will: From Metaphysical Freedom to Normative Functionalism*, 36 J. AM. ACAD. PSYCHIATRY & L. 16, 16 (2008) ("It is widely assumed that people have a relatively unhindered capacity to make choices and to decide what they will do."); Sousa, *supra* note 36, at 11 (recounting study results that a majority of respondents thought a universe of indeterministic choice was most like ours).

145. *See* Farah, *supra* note 69, at 38.

146. *Id.*

147. PAUL M. SNIDERMAN ET AL., REASONING AND CHOICE: EXPLORATIONS IN POLITICAL PSYCHOLOGY 207 (1991).

148. *Id.*

149. The idea that an act cannot entail blame without a guilty mind is often said to be the cornerstone of modern criminal law. *See* Duncan v. State, 26 Tenn. (7 Hum.) 148, 150 ("It is a sacred principle of criminal jurisprudence that the intention to commit the crime is of the essence of the crime . . . ."); *see generally* Albert Levitt, *The Origins of the Doctrine of Mens Rea*, 17 U. ILL. L. REV. 117 *passim* (1922) (describing the influence of ecclesiastical law on

undoing the very heart of how law operates and the connection between law and the people it governs.

But neurolaw's promise to reveal why people think and behave as they do is inescapably built on the idea that people are not agents as the law traditionally views them. Instead, they are guided almost entirely by determined and unconscious chemical cascades which exert irresistible control over an agent's thinking and behavior. What is needed is not blame and punishment but therapeutic interventions formulated by science which can control and mitigate future harm. A neurolaw informed criminal justice policy sees criminal behavior as the result of abnormal brain processes, not the exercise of willful choice by free agents. What matters is not whether those agents knew that their behavior was wrong, but whether they can control it in the future. And understanding whether an agent has the ability to do so rests entirely under the expertise of science, not judges or juries. Neurolaw seeks not just to inform the law, but also to impose its supremacy in crafting policy. It begins this endeavor by undoing the distinction between the criminal and civil code.

A. MAKING THE CRIMINAL, CIVIL

Criminal law proscribes conduct. Yet most criminal statutes begin their proscriptions not with behavior but with mental states.[150] Modern criminal law goes to great lengths to differentiate between those who intend their illicit conduct from those who fail to exercise prudence even if the conduct produces the same result. The difference between murder and manslaughter generally falls not on the end result of the agent's behavior but her mental state at the time of the offense. Thus, criminal law follows ordinary intuition which hold that death caused by the intentional or malicious conduct of a defendant should be treated very differently than when it is caused by reckless or careless behavior. It is from this common intuition that mercy and mitigation flow.[151] Even when agents

---

modern Western legal codes which emphasized the mental state in determining blameworthiness).

150. *Cf.* JEROME HALL, GENERAL PRINCIPLES OF CRIMINAL LAW 70 (2d ed. 1960) ("The principle of mens rea is the ultimate evaluation of criminal conduct and, because of that, it is deeply involved in theories of punishment.").

151. *See* O. Carter Snead, *Neuroimaging and the "Complexity" of Capital Punishment,* 82 N.Y.U. L. REV. 1265, 1309–12(2007) (describing the nexus between retribution, mitigation, and intuition).

engage in intentional conduct which violates the law, their related blameworthiness is weighed against facts which might suggest reduced culpability.[152] The affirmative defenses of insanity, extreme emotional disturbance, or duress exist because the law is grounded in the notion that for most proscriptions, mental states matter in determining culpability.[153]

But it was not always this way. The view that a defendant's mental state mattered in determining culpability is an outgrowth of important developments in enlightened society. Before the Twelfth Century, strict liability was the regular course for criminal law.[154] The view that extenuating circumstances should matter in determining not only a defendant's culpability but his punishment as well was entirely foreign. People were held strictly accountable and punished for their behavior alone.[155] Whether someone intended to commit a crime or caused the prohibited conduct by recklessness or negligence was irrelevant to determinations of culpability and dispensing of punishments.[156] The pre-twelfth criminal justice system operated efficiently and brutally: Trials were quick and dispensation of punishment quicker.[157]

Succeeding years brought nuance to the criminal law.[158] Many critics of the current criminal code quite rightly lament its breadth and depth.[159] Indeed, many behaviors are regulated and outlawed. And many of those are outlawed repeatedly through numerous layers of jurisdictions. But what is often neglected in this discussion is that the birth of this growth was partly due to the emergence of the doctrine of mens rea. That is, criminal law began its enlargement under the elementary premise that it mattered *more* what an agent intended than the

152. *See* Farah, *supra* note 69, at 38.

153. *Id.*

154. *See* Francis Bowes Sayre, *Mens Rea*, 45 HARV. L. REV. 974, 977 (1932) ("[F]or whatever the law in action may have been, [fragments of early English law] show that at least in the recorded law prior to the twelfth century, a criminal intent was not recognized as an indispensable requisite for criminality.").

155. *Id.* at 977–78 n.9.

156. *Id.*

157. *See, e.g*, *id.* at 976 (discussing trial by battle).

158. *See id.* at 1004–17.

159. *See* William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 507–08, 519–23 (2001).

actual harm caused by his conduct.[160] Whether an agent should be punished severely or excused for causing the death of another is determined not by the behavior which resulted in the criminal conduct but by what he (or a reasonable person in the same situation) knew and intended.[161] And while many scholars have questioned the nature of the reasonable person standard, those critiques largely agree that modern law gives credence to an agent's mental states.[162] Thus, justified homicide results in no punishment or blame because the law freely accepts the idea that it is permissible to kill another when under threat of immediate death from an aggressor.

And to be free of blame is what results in freedom from punishment. The law holds the agent who justifiably kills outside the purview of punishment because to do otherwise would offend ordinary intuitions of blameworthiness.[163] Likewise, the law views predatory behavior as especially egregious because ordinary intuitions of justice suggest that agents who premeditate and carry out calculated crimes against innocent victims represent the pinnacle of blameworthy conduct which the law finds so offensive.[164] It is offensive and punished severely not because the resulting harm is the greatest in any consequential sense, but simply because it so deeply trespasses against the shared intuitions of proper conduct venerated by most people. And those intuitions are shared not just provincially, but across cultures and generations.[165] Despite all of our individual differences, most people agree that intentional unlawful conduct is worse than negligent behavior irrespective of the resulting harm.[166]

These shared intuitions of justice can reveal much about the way people conceptualize behavior, blame, and punishment. A common critique suggests that in thinking about crime, people make judgments based on emotion.[167] Indeed, several

160. *See* Sayre, *supra* note 154, at 988–94.

161. *Id.*

162. *See, e.g.,* MAYO MORAN, RETHINKING THE REASONABLE PERSON 5–17 (2003).

163. *See* Robinson & Darley, *supra* note 2, at 10.

164. *Id.*

165. *Id.*

166. For a review of empirical studies examining intuitions of justice across demographics *see* Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice,* 91 MINN. L. REV. 1829 (2007).

167. Usually the emotion indentified in driving judgments about crime is fear. *See* DAVID GARLAND, THE CULTURE OF CONTROL: CRIME AND SOCIAL

cognitive neuroscience studies seem to suggest this is true.[168] Yet the subtle premise among such critiques is that this is inherently a bad thing and the legal process should do what it can to minimize it.[169] And perhaps in some ways it should. But not much attention is paid to how such practices might destroy the common bond between the public and the criminal code which proscribes their conduct. If shared intuitions of justice are indeed mutual across all cultures then they serve as a touchstone of something that is uniquely emblematic of our humanity. At the least, these intuitions tell us what it means to hold someone blameworthy and not just responsible.

Indeed, what makes the criminal law different from the civil code is the unique way it blames and punishes. When a tortfeasor is held liable, we blame her for her wrongdoing. We may even feel strong condemnation for her actions. But we do not make determinations about her character or value to society in the same way those judgments are exercised in criminal cases. Moral blameworthiness is the exclusive domain

---

ORDER IN CONTEMPORARY SOCIETY 10 (2001) (discussing the invocation of fear in crime control policy during the 1970s). *But see* Emily Gray et al., *Reassessing the Fear of Crime*, 5 EUR. J. CRIMINOLOGY 363, 373–77 (2008) (questioning the scope of that fear on methodological grounds). Vengeance is also commonly identified as being expressed in the criminal law. *See* Robert C. Solomon, *Justice v. Vengeance: On Law and the Satisfaction of Emotion*, *in* THE PASSIONS OF LAW 123 *passim* (Susan A. Bandes ed., 1999). For an overview of psychological research into emotions, see Neal Feigenson & Jaihyun Park, *Emotions and Attributions of Legal Responsibility and Blame: A Research Review*, 30 L. & HUM. BEHAV. 143 (2006).

168. *See* Joshua W. Buckholtz et al. *The Neural Correlates of Third-Party Punishment* 60 NEURON 930, 935 (2008); Ming Hsu et al., *The Right and the Good: Distributive Justice and Neural Encoding of Equity and Efficiency,* 320 SCI. 1092, 1095 (2008).

169. These arguments are often made based on examination of victim impact statements and crime scene photographs during trial. *See, e.g.,* Jessica M. Salerno & Bette L. Bottoms, *Emotional Evidence and Jurors' Judgments: The Promise of Neuroscience for Informing Psychology and Law*, 27 BEHAV. SCI. & L. 274, 275–82 (2009) (reviewing studies of "emotional evidence" and arguing for a greater understanding of the role that emotion plays in jury decision-making). *But see* Paul G. Cassell, *In Defense of Victim Impact Statements*, 6 OHIO ST. J. CRIM L. 611 (2009) (discussing empirical evidence which suggests no significant association between victim impact statements and sentencing outcomes); Theodore Eisenberg et al., *Victim Characteristics and Victim Impact Evidence in South Carolina Capital Cases*, 88 CORNELL L. REV. 308, 309–10, 341 (2003) (finding no significant association between victim impact statements and sentencing outcomes in over 200 cases).

of the criminal law.[170] It judges and condemns people for behavior which so deeply intrudes upon common intuitions of what people view as fundamentally necessary and right in an ordered society. It is hardly surprising, then, that criminal law traditionally reserves its most severe punishment for crimes which offend this intuitional (and institutional) social order: murder, rape, assault, and theft. It also is unsurprising that society seems rather undaunted about whether these punishments are consequentially effective at deterring crime. Under this lens, criminal law is not about consequential gains, but it is entirely rationale in a folk psychological sense.

But those who view behavior from a cognitive neuroscience perspective are hardly impressed with these shared intuitions gleaned from generations. Instead, blame represents a vestige of our primordial past—the all too frequent override of our primitive and emotional brain over our recently acquired cortex.[171] They see our criminal justice system as foolishly wed to intuitions which are prone to biases and imprudently concerned with an agent's intent.[172] Indeed, the very notion that criminal behavior should be viewed as a criminal matter whatsoever makes little sense. Rather, crime is a sign of underlying brain pathology.[173] Those who steal have impulsive brains;[174] those who inflict pain and physical misery unto others are said to lack an empathic brain.[175] Crime itself is a sign of epidemic mental disorder throughout the land.

Punishment is irrational from this perspective. If agents are largely unable to control their conduct because of the determined nature of their brains, then it makes little sense to punish people for their conduct. Since punishment begins with social condemnation, it wrongly assumes that the convicted

---

170. *See* Paul H. Robinson, *The Criminal-Civil Distinction and the Utility of Desert*, 76 B. U. L. REV. 201, 202 (1996) ("Central to lay person's view is that criminal sanctions signal moral condemnation, while civil penalties do not.").

171. *See* Johannes Haushofer & Ernst Fehr, *You Shouldn't Have: Your Brain on Others Crimes*, 60 NEURON 738,739 (2008).

172. *See* Greene & Cohen, *supra* note 3, at 1776.

173. *See* Adrian Raine, *From Genes to Brain to Antisocial Behavior*, 17 CURRENT DIRECTIONS IN PSYCHOL. SCI. 323, 323 (2008) ("The key concept highlighted in this review is that specific genes result in structural and functional brain alterations that, in turn, predispose to antisocial behavior.").

174. *See generally id.* at 324 (explaining that those who commit violent crime have impulsive brains).

175. *See* Alison Abbott, *Abnormal Neuroscience: Scanning Psychopaths*, 450 NATURE 942, 942 (2007).

agent exercised control over his conduct. Instead of punishment, cognitive neuroscience desires to deploy a panoply of therapeutic interventions aimed at ameliorating future harm.[176] Thus, instead of looking back at an agent's wrongful conduct as the meter of his deserved punishment, cognitive neuroscience looks exclusively forward at an agent's risk of future dangerousness.[177] And while this criminal justice policy holds some luster because it measures deprivations of liberty entirely from a consequential perspective, it also accomplishes two secret ambitions of the cognitive neuroscience movement. First, it severs the measurement of punishment from ordinary intuitions of justice and fairness Second, it conveys construction of crime policy from a legal framework to one engineered by science.

The amputation of punishment from the public's intuitions is the grand hope which rides near the crest of neurolaw's ambition. Few articles which discuss neuroscience and punishment fail to mention what is perceived as the malignant pathology within crime policy these days: the rise of retribution.[178] And the blame for this rebirth of desert falls squarely on the influence that common intuitions of justice play in crafting crime policy.[179] Critics contend that our modern criminal justice system is overwrought with placating the public's demand for desert.[180] Instead, they argue, we should entrust a forward-focused system, guided by experts who will dispassionately gauge confinement based on an offender's propensities. After all, propensities can be ameliorated and managed, but what has happened in the past cannot be undone.

While the critics may be right, they are short-sighted: mitigation and mercy are direct descendants of desert.[181] Mercy

176. Although few have argued directly for this result, it is a presumed goal given the foreseen influence of neuroscience with the criminal justice system and the ubiquity of the technology. *Cf.* Greene & Cohen, *supra* note 3, at 1781 ("At some further point this sort of brainware may be very widespread, with a high-resolution brain scanner in every classroom. People may grow up completely used to the idea that every decision is a thoroughly mechanical process. . . .").

177. *Id.* at 1776.

178. *Id.* at 1776–78.

179. *Id.*

180. *Id.*

181. *See* Snead, *supra* note 151, at 1319–22.

does not flow from an agent's future propensities, but from her past circumstances which suggest she is less blameworthy. We feel sorry for the defendant who was repeatedly abused as a child, even if that abuse suggests a heightened risk for future harm to the public.[182] If ordinary intuitions are indeed based on emotions, those emotions forgive as well as condemn. It may be the case that modern society does not enough of the former and too much of the latter, but employing more mercy may very well entail opening the emotional spigot of intuitions instead of welding them shut entirely.[183]

And while focusing on an offender's future propensities may appear more sensible than beckoning blame, it is dubious whether this will result in a more humane criminal justice system. Future propensities are notoriously difficult to externally manage. The work in criminal psychology has plainly demonstrated that two factors are largely responsible for recidivism across a diverse population of offenders. The first is alcohol and drug abuse; the second are dispositional traits that run to the core of an agent's personality.[184] Both are exceptionally recalcitrant to change and easily evasive of external monitoring even in therapeutic settings. But that has not prevented a host of legal prescriptions based on the therapeutic lens from being implemented for offenders deemed mentally unsound. And those legal prescriptions are hardly favorable for criminal defendants.

For years, most sex offenders were punished and sentenced to prison. In 1996, the Supreme Court decided the case *Kansas v. Henricks,*[185] which upheld a statute authorizing the civil commitment of sex offenders after they had served their prison sentences if they suffered from a mental abnormality. The statute merely required that a person subject to the statute have a condition which affected his "emotional or volitional

182. *See* Cathy Spatz Widom, *Child Abuse, Neglect, and Violent Criminal Behavior,* 27 CRIMINOLOGY 251, 260–63 (1989) (reporting a significant association between child abuse and arrests for violence during adulthood among a large sample of children).

183. The rise of the administrative state which is very isolated from common intuitions probably only hastens that trend. *See* Rachel E. Barkow, *The Ascent of the Administrative State and the Demise of Mercy*, 121 HARV. L. REV. 1332, 1362–63 (2008).

184. See D.A. ANDREWS & JAMES BONTA, THE PSYCHOLOGY OF CRIMINAL CONDUCT (4th Ed. 2006) for a methodical overview of the identified risk factors for criminal recidivism.

185. Kansas v. Hendricks, 521 U.S. 346, 360, 371 (1997).

capacity"[186] in a manner which predisposed him to commit sexually violent conduct. Because the aim of the statute was therapeutic and not punitive, Constitutional protections afforded to defendants were unavailable to Henricks. The prolonged incapacitation of people like Henricks is made possible by the very fact that they are within the civil mental health system and not under the regime of criminal punishment with its enumerated Constitutional protections.

It may be difficult to feel much sympathy for people like Henricks, but his indefinite confinement is only possible because he is deemed to have a propensity for unlawful conduct caused by an amorphous mental abnormality that is considered amenable to therapeutic treatment. His predilection for pedophilia is a mental abnormality not just because the aim of his pursuits is illicit, but because he is said to lack control over his behavior. This dual determination that the content of his desires and inability to control them are a *mental* abnormality may have some folk psychological grounding. But the architect of the mentally disordered mindset prevalent in our culture is science itself. After all, it is the official diagnostic manual of psychiatry which holds seventeen different mental disorders based on sexual conduct.[187]

Inasmuch as cognitive neuroscience promises to reveal how people think, its message has also been that behavior is routinely outside of the control of the agent.[188] The reason why implicit biases matter is not that people may have them, but that they act on them.[189] What neurolaw openly admits is what implicitly grants it power: that it can reveal how the mind operates and those revelations mean that science can craft

186. KAN. STAT. ANN. § 59-29a01 (1994).

187. DSM-IV REVISION, *supra* note 59, at 535–82.

188. *See generally* WEGNER, *supra* note 43, at 1–28 (arguing that the experience of consciously willing an action is not a direction indication that the conscious thought has caused the action—that it is actually a complex illusion).

189. The concept of implicit biases has received much scholarly attention, due in large part, to results of numerous studies utilizing the Implicit Association Test. For an overview, *see* Anthony G. Greenwaldt & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations,* 94 CAL. L. REV. 945, 951 (2006). *But see* Hart Blanton et al., *Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT*, 94 J. OF APPLIED PSYCHOL. 567, 580–81(2009) (heavily criticizing the psychometric properties of the Implicit Association Test).

policy better than the law.[190] That is so because the mind executes behavior and it is conduct that the law regulates. In a determined world where minds occupy harmful intentions unknown even to those who possess them, the promise of science is its ability to discern with reasonable definitiveness how and when people will act on those intentions. This ineluctably places science in a better position in formulating crime policy when the only legitimate distributive principles of punishment are consequential ones.

The coming transformation of our criminal justice system promised by the cognitive neuroscientists thus rests on a deceptively simple tripartite set of premises which has a profound implication. Quite audaciously, they assert that the human mind is a determined entity *whose processes can be entirely revealed by the field of neuroscience*; the determined nature of the mind means that concepts such as intention are illusory; and deserved punishment is illogical. The implications of these positions, however, bluntly challenge the nature and justification for a criminal justice system itself. If what differentiates the criminal justice system from civil regulation is its ability to punish for the sake of moral condemnation, then under the tripartite model of cognitive neuroscience the criminal justice system should dissolve as "the idea of distinguishing the truly, deeply guilty from those who are merely victims of neuronal circumstances" will indeed be pointless.[191] After all, what could be considered criminal without a guilty mind?[192]

B. RISE OF THE CONTROL TESTS

In most cases, it is axiomatic that to blame an agent of criminal conduct is to accuse her of possessing a guilty mind at the time of the offense. This link between blame and mind is a foundational bridge between law and the common intuitions of the people it governs. That is, the law holds to a strong account of human agency harmonious with how most people believe the

190. *See* Greene & Cohen, *supra* note 3, at 1781 (describing how neuroscience will play a central role in crime policy by making "detailed predictions about how the mechanical processes [of the mind] work").

191. *See id.*

192. *See* Levitt, *supra* note 149, at 117 ("Nearly all the courts say that there can be no criminal act unless there has been a criminal intent; that a guilty mind must be present or no crime is committed.").

mind operates.[193] For the law, brains do not commit to contracts, become married, or commit crimes—people do.[194] Thus, people are accountable for their behaviors and the law can legitimately punish people for engaging in behaviors which the law dictates run afoul of the established code. That code is in many respects a social contract between the people and its government. [195] It is also a contract which provides assurances to individual citizens that they are expected to act lawfully irrespective of their personal propensities or liabilities towards unlawfulness. In turn, fellow citizens can expect the same from others. It is in this fashion that people can reasonably assume others will refrain from untoward acts of harm despite their impulse or robust desire to do otherwise. In sum, choices matter for personal responsibility under the law and in our society as well.[196]

But the law has also long recognized that under certain circumstances, people are not responsible for their behavior. Substantive criminal law in most jurisdictions provides that a person is not guilty of a crime if he did not know or appreciate the wrongfulness of his actions due to a mental illness or defect. Almost all reasonable scientists these days hold to the

---

193. As Chief Justice Biggs put it:

> The concept of mens rea, guilty mind, is based on the assumption that a person has a capacity to control his behavior and to choose between alternative courses of conduct. This assumption, though not unquestioned by theologians, philosophers and scientists, is necessary to the maintenance and administration of social controls. It is only through this assumption that society has found it possible to impose duties and create liabilities designed to safeguard persons and property.

United States v. Currens, 290 F.2d 751, 773 (3d Cir. 1961); *see also* Charles C. Steward Mach. Co. v. Davis, 301 U.S. 548, 590 (1936) ("[T]he law has been guided by a robust common sense which assumes the freedom of the will as a working hypothesis in the solution of its problems."); Speranza, *supra* note 34, at 125 ("Law, on the other hand, stands pre-eminently for the freedom of the will.").

194. *See* Stephen J. Morse, *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*, 3 OHIO ST. J. CRIM. L. 397, 397 (2006).

195. *Cf.* Gregg Cartage & Storage Co. v. United States, 316 U.S. 74, 79–80 (1942) ("[T]he practical business of government and administration of the law is obliged to proceed on more or less rough and ready judgments based on the assumption that mature and rational persons are in control of their own conduct.").

196. *See* 1 GEORGE P. FLETCHER, THE GRAMMAR OF CRIMINAL LAW 266 (2007) ("The capacity for executing choices provides the foundations for responsibility in criminal law.").

biological hypothesis of mental illness.[197] Those illnesses most associated with insanity claims—psychotic illnesses—are presumed to originate in the brain and are principally involved with much of the behavior concerning claims of reduced culpability. Indeed, some of the most compelling evidence from neuroscience that bears on brains and behavior comes from studies demonstrating severe and enduring global deficits in the brains of people afflicted with psychotic illnesses.[198] As commonly understood, the person is excused because of his impaired brain. That is, his afflicted brain is believed to obscure his true mind. But cognitive neuroscientists and their supporters claim that this distinction between brains, minds, and people is superfluous: all that matters is the brain.[199] If that is true, then excusing agents because of their brains seems absurd. Brains cannot excuse the mind because they are one and the same.

But what cognitive neuroscience supporters are really suggesting is not that excuse is illogical, but that that the legal emphasis on a defendant's cognitive state is misplaced. In their view, an agent's self-awareness is deceptive because they are entertained by a stout illusion of free will. Ascertaining whether an agent possessed a guilty mind at some point in the past is not only difficult but pointless. Determinations of whether an agent knew his behavior was wrong conflate knowledge with intention; to say an agent understood that his behavior was unlawful and acted anyway wrongly implies that he could have intended otherwise.

Since cognitive neuroscience views intentions as illusory,

---

197. *See* Carl I. Cohen, *The Biomedicalization of Psychiatry: A Critical Overview*, 29 COMTY. MENTAL HEALTH J. 509, 509–10 (1993) (noting that biological psychiatry is now the dominant version of psychiatry); *see also*, Samuel H. Barondes, *The Biological Approach to Psychiatry: History and Prospects*, 10 J. NEUROSCIENCE 1707, 1708–09 (1990) (discussing the rise of biological psychiatry).

198. *See* Christos Pantelis et al., *Structural Brain Imaging Evidence for Multiple Pathological Processes at Different Stages of Brain Development in Schizophrenia*, 31 SCHIZOPHRENIA BULL. 672 *passim* (2005) (reviewing the wealth of studies reporting numerous and significant abnormalities in brains of those afflicted with schizophrenia including progressive changes that occur during the earliest stages of the disease, often before the initiation of pharmacotherapy); *see also* Paul J. Harris & David A. Lewis, *Neuropathology of Schizophrenia*, *in* SCHIZOPHRENIA 310, 310–25 (Steven R. Hirsch & Daniel Weinberger eds., 2003) for a general discussion on the various brain abnormalities associated with schizophrenia.

199. *See* Greene & Cohen, *supra* note 3, at 1779.

what matters are the biological and genetic propensities of the agent which are exercised from his brain.[200] And while these properties are impinged by the environmental forces encountered by the agent during everyday life, the neurolaw view holds the agent as hapless in avoiding any of the untoward effects of those interactions. These behavioral propensities lodged deep within the brain are thought to result directly in conduct irrespective of an agent's desires or wishes otherwise. Whatever self-awareness or willful movement actually entails, it's domain over the biological circuitry of behavior is trivial.

What this entails for criminal responsibility at the margins is a move towards control tests in excuse and diminished capacity defenses. That is, neurolaw asserts that what matters in determining an agent's culpability is her volition and not her intentions or knowledge. By claiming science has illuminated that people operate in a manner quite divergent from their folk psychological beliefs, neurolaw hopes to bring biology to the foreground in forming social and legal norms. And because the disparity between common intuitions and cognitive neuroscience are their greatest in the realm of responsibility, neurolaw has focused its attention at influencing criminal law doctrine where accounts of responsibility are most salient.

Supporters claim three closely related benefits from this neuro-jurisprudence model of criminal law. First, it removes blame from determinations of guilt and calculations of punishment. Because blame is viewed as a farrago of expressive judgments unwisely tethered to human passions, its impact on criminal adjudications can be avoided.[201] Second, neuroscience will advance a preventive model of criminal justice aimed squarely at avoiding future harms.[202] And third, neuroscience, aided by other fields of psychological science, can uncouple culpability determinations from the arm of criminal sanctions.[203]

---

200. *Id.* at 1781. ("[Y]our brain serves as a bottleneck for all the forces spread throughout the universe of your past that affect who you are and what you do.").

201. *See id.* at 1782.

202. *Id*. at 1783.

203. *See* Slobogin, *supra* note 9, at 165 (stating that "[t]he punishment model of the criminal law is currently threatened by the newly popular prevention model of intervention, one that is based on predictions of risk

All of these claimed benefits rest on the larger premise proffered by cognitive neuroscience which holds that agent's self-awareness greatly inflates her true ability to control her conduct. Behavior generates from discharges of neuronal output which are largely (if not entirely) beyond the influence of consciousness. As a consequence, excuse should rest solely with determinations of volition.

Legal formulations of excuse vary over jurisdictions and have unquestionably been influenced by cultural perceptions of the mind and mental illness.[204] Yet quintessentially the doctrine holds that a person is not responsible for a criminal conduct if during the commission of the crime, due to a mental illness or defect, he lacks substantial capacity to either appreciate the criminality of his conduct or conform his conduct to the requirements of the law.[205] This standard test holds that excuse is warranted when a defendant lacks the awareness to know right from wrong or is unable to control himself.[206] And while the time and venue have favored one part of this test over the other, the cognitive prong has resolutely remained intact over the years.

But cognitive neuroscience promises fundamental change to the excuse doctrine. While it confidently articulates an ability to discern how people think, it also suggests that individual awareness plays a minimal role in the course of day to day conduct. Instead, behavior is mostly an autonomous process whereby people mechanically respond to stimuli presented to them.[207] Concepts deeply engrained in criminal law, such as intent, are considered remnants of the folk psychological belief required by society that obeying the law necessitates thinking and choosing by an agent. That this view might suggest a radical overhaul of criminal liability in its entirety has not gone without notice among its proponents—indeed much of the neurolaw literature gleefully anticipates it. But institutional change tends to come incrementally. The emerging dominance of control tests in culpability assessments is the first brick removed from the

---

uncabined by culpability assessments.").

204. *See* MICHAEL L. PERLIN, THE JURISPRUDENCE OF THE INSANITY DEFENSE (1993).

205. *See, e.g.,* MODEL PENAL CODE § 4.01 (1968).

206. *Id.*

207. *See* ALBERT BANDURA, SOCIAL FOUNDATIONS OF THOUGHT AND ACTION: A SOCIAL COGNITIVE THEORY 12 (1986).

pillar of mens rea in criminal law.

And while institutional change is a necessary force within law, science stands apart from political and other authorities who impinge legal doctrine because it is seen as empirical and objective. Arguments favoring the demise of desert in favor of consequential sanctions can be appealing for a variety of reasons. But those which rely heavily on neuroscience entail prediction of future behavior as the linchpin of their reforms. Instead of utilizing common intuitions of guilt as the meter for deserved punishment, those who propose a neuro-person model of culpability place their faith in the ability of the behavioral sciences to accurately predict future behavior. Control tests satisfy this reform because they imply a persistent impairment of behavioral capacity.

While prediction of behavior has endured its share of heavy criticism over the years, the rise of a new method is enjoying wide acceptance from mental health professionals and the courts. Actuarial risk assessment uses crude analysis of historical data to predict future behavior—ranging from sexual recidivism to violence.[208] These methods are already broadly employed by experts in all stages of civil and criminal adjudications where risk of unlawful or undesirable conduct is of concern.[209] Due in large measure to the legal conversion of acts of rape and pedophilia from crimes to mental abnormalities,[210] cases turning on expert assessment of future dangerousness are legion.[211] Consideration of future dangerousness by way of these behavioral assessments dominates sentencing in capital cases as well.[212] Sentencing alternatives, including the heralded drug and mental health

208. *See* John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 408–09 (2006).

209. *Id*. at 408-09 nn.71-72.

210. The behaviors associated with these crimes are now deemed mental abnormalities as well as criminal conduct under the law. *See* Erickson, *supra* note 56, at 72–73.

211. *See* BERNARD E. HARCOURT, AGAINST PREDICTION 39–108 (2007); Monahan, *supra* note 208, at 396–404; *see also* David DeMatteo & John F. Edens, *The Role and Relevance of the Psychopathy Checklist-Revised in Court: A Case Law Survey of U.S. Courts (1991-2004)*, 12 PSYCHOL., PUB. POL'Y, & L. 214, 214 (2006) (noting the prevalence of admitting an instrument for measuring psychopathy in civil and criminal cases).

212. *See* Snead supra, note 151, at 1322–24 (2007).

courts, are premised on the idea that interventional behavioral treatments will reduce future illegal behavior in predictable ways—and that those who may benefit from these interventions can be determined and selected from the general pool of offenders.[213] The marriage between behavioral prediction and interventional therapeutics is strong and growing stronger by the day.

What is often neglected in discussions about this new science of prediction, however, is how closely the science mirrors common intuitions. The empirical literature within the behavioral sciences strongly suggests that the largest risk factors for future unlawful behavior are ones found in lay judgments about punishment and risk. Prior bad acts, chronic drug abuse, and an offender's penchant for violating social norms weigh heavily in favor of recidivism in both camps. And while science can claim that its approach is methodically and analytically superior, it has yet to demonstrate how its approach generalizes to superior understandings about population trends. That is, science is on par with lay intuitions about individual recidivism factors, but it remains a mystery why crime rates rise and fall over time despite all of the sophisticated empirical analyses employed over the years—just as most lay intuitions seem dumbfounded to explain these trends.

The prediction model, however, openly favors experts. It desires distributive punishment as an enterprise that has a strong fidelity to incapacitation and the prevention of future harms. And within this model, someone must judge and determine who is dangerous and should be detained and who is harmless and should go free. To accomplish this task, the prediction model must invariably engage in the precarious side of what lay intuitions do in ascertaining risk: grouping people by their shared traits. Psychological science has long demonstrated that people have the tendency to make judgments about others—particularly when judging risk—by using heuristic cues about individuals that imply membership to a larger group.[214] Whether that means the individual is left-handed and the person judging him implicitly views all left-handed people as a passive bores entertained by math and

213. *See* Greg Berman & John Feinblatt, *Problem-Solving Courts: A Brief Primer*, 23 LAW & POL'Y 125, 131–135 (2001).

214. *See* Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1124 (1974).

crossword puzzles or the individual retains membership to a certain ethnicity and the person views that ethnicity with suspicion or contempt, unconscious cues seem to matter in how people make day to day judgments about risk.

To make predictions about a person's risk of future behavior means placing that person within a group with known risk factors. No one seriously entertains the idea that a prediction model based on neuroscience would engage in overt racism by pooling race as a risk factor. The difference between lay intuitions which group people based on heuristics and a prediction model assembled by experts would be that the latter would not be influenced by overt prejudices. But inescapably, that model would hold manifestly genetic factors as important and thus group offenders together because of that fact. Indeed, this is already being done. One of the most prevalent sex offender recidivism assessment tools employed by experts weights gender as a significant risk factor[215]—as it should.

The harm is not that future prevention models would routinely employ genetic factors such as gender in determinations of future dangerousness. They would be foolish not to do so. But wedding cognitive neuroscience and the prevention model risks classifying brains based on what is not manifestly evident to most people. To say that an offender poses a risk of future harm because of deficient neurotransmitter production[216] or slight delays in pressing keys when viewing illicit images[217] removes these determinations from any grounding in common intuitions. That those who support the prediction model suggest that people outside of the traditional purview of the criminal justice system might also be

215. Known as the Static-99, this actuarial scale uses gender as one of the ten items it evaluates. *See* Static-99 Clearinghouse, http://www.static99.org/ (last visited Nov. 6, 2009).

216. A proposed link between deficit production of a neurotransmitter known as serotonin and violence was established less than a decade ago. *See* Avshalom Caspi et al., *Role of Genotype in the Cycle of Violence in Maltreated Children*, 297 SCI. 851, 853 (2002). *But see* Niel Risch et al., *Interaction Between the Serotonin Transporter Gene (5-HTTLPR), Stressful Life Events, and Risk of Depression: A Meta-Analysis*, 301 J. AM. MED. ASS'N 2462, 2468 (showing no link between the proposed serotonin gene and psychological disorder).

217. *See* Gillian Smith & Lane Fischer, *Assessment of Juvenile Sexual Offenders: Reliability and Validity of the Abel Assessment,* 11 SEXUAL ABUSE 207 *passim* (1999) for a review of this controversial, yet widely used psychological test.

within the scope of such a model suggests the reach would be massive. Rather than reverse the breadth and depth of the current criminal justice system, the prediction model seems all too likely to expand it greatly.

C. ABOLITION OF AGENCY

Most people believe that they have an ability to impose their desires and intentions in the world. They also believe that others, absent severe mental or neurological disease, do the same. In common sense terms, as well, some people have more difficulty controlling their conduct than others.[218] Some can indulge in alcohol to moderation while others cannot, in part, because of their biology. To be born means to be born with biological attributes and liabilities.

The law embodies this view of humanity because it is fundamentally a human enterprise. The law of legislatures and courts is not the law of physics or chemistry. It does not exist independent of the judgments of those who are subject to its rules. Rather, it exists because those who are subject to its rules value certain beliefs and ideals which the law incorporates into its code. The force of law is bolstered or subdued by the bind of humanity which is under its rule.

Justice means different things to different people, but few would hold that letting the guilty go free or punishing the innocent are just outcomes. These shared beliefs about justice reveal important aspects of the law which are necessary for people to view it as legitimate and fair. Nowhere are these shared beliefs stronger than in conduct deemed criminal—they go to the core of what it means for law to do justice.[219]

These shared intuitions, then, are vital to the criminal law's operation. Irrespective of whether they are imprudently influenced by emotions or prone to manipulation by cunning

---

218. *See* Stephen J. Morse, *Against Control Tests for Criminal Responsibility*, *in* CRIMINAL LAW CONVERSATIONS 1 (Paul H. Robinson et al. eds., 2009).

219. *See* Paul H. Robinson et al., *The Origins of Shared Intuitions of Justice,* 60 VAND. L. REV. 1633, 1634–38 (2007) (demonstrating a consensus on the core intuitions of justice even across demographics and cultures). This core consensus occurs during early development. *See* Jonathan Haidt et al., *Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?*, 65 J. PERSONALITY & SOC. PSYCHOL. 613, 617, 621–22 (1993) (observing that children ten to twelve years old across the cultures investigated think that pushing another child off a swing should be punished and that this behavior would be wrong in other countries as well).

persuasion,[220] they lie at the heart of what it means to condemn others and justify punishment. We punish and condemn because we blame people for behaving wrongly. And that blame is not just a statement that someone has transgressed against the law, but is a proclamation that they have violated deeply held social norms important to most people.[221]

The mode of neurolaw is one divorced from this understanding. Its explanations are geared toward undoing conventional models of crime and punishment in favor of one built on mechanistic accounts of behavior. Aided by technology that few understand, neurolaw's draw is its promise to unravel the mysteries of the mind and demonstrate that thinking and behavior are predictable given its mechanical and determined nature—that the mind is not that great of a mystery after all.

So, too, does the prediction model seek in unraveling the influence of intuitions from judgments of guilt and punishment. As a forward-focused enterprise, prediction is liberated from determinations of blameworthiness. Instead, its goal is cabined within the expertise of science in judging future harm based on the view that incapacitation as the central pillar of crime policy distributes justice more fairly and rationally.

Both neurolaw's explanation of behavior and the prediction model's goal of moving sentencing determinations into the realm of prevention diminish the concept of the guilty mind in criminal law. Neurolaw claims awareness and intention are minefields of illusions; prediction models view the past intentions of an agent as irrelevant. Uncoupling guilt from criminal sanctions is considered both principled and prudential: the guilty mind requirement is extraneous and pointless; it also stands in the way of a criminal justice system based on prevention of future harm. And prevention of future harm necessarily means employing legal sanctions before the occurrence of unlawful conduct—before the criminal act arises. After all, that is what preventive detention is all about.

The breadth and depth of that system of criminal justice will be massive. The real danger will not be that neurolaw will

---

220. The argument against intuitional judgments usually falls on the notion that they are emotionally laden and irrational. *See* JESSE J. PRINZ, THE EMOTIONAL CONSTRUCTION OF MORALS (2007).

221. *Cf.* Jeffery J. Rachlinski, *The Limits of Social Norms*, 74 CHI.-KENT L. REV. 1537, 1539–42 (2000) (discussing the difficulties in using social norms for public policy).

completely undo criminal law but that it will retain the vast power of criminal law to sanction behavior while removing personal responsibility from the legal presumption that unlawful conduct deserves to be punished. Unlawful behavior will be seen as indicative of a sick mind with its agent hopelessly along for the ride. Rational crime policy will entail therapeutic interventions instead of punishment. Those interventions will have all of the callings of the rehabilitative ideal but will rest squarely on calculated determinations of future harm. And there is little reason to believe that those interventions will be any less intrusive or the determinations any more liberating than the current system. Indeed, there is every reason to believe they will fail to meet those objectives.

Understanding the consequences of neurolaw's vision for the criminal justice system begins with appreciating its ambition. As a movement which considers the mind as entirely reducible to explanatory models of neuronal discharge, it steadfastly maintains that all human mentation will be revealed through technology. The plain purpose of this knowledge will be to grant the natural sciences authority in defining and explaining not just human behavior but the very essence of humanity. Such authority will be difficult for the law to resist since it is fundamentally rooted in regulating human behavior. That authority will play a larger role in formulating crime policy, and in so doing, will ineluctably lead towards polices which emphasize behavioral prediction and incapacitation.

But this is also a movement that views folk psychological ideas such as the guilty mind with deep suspicion. As a consequence, the demarcation between what the law considers criminal and civil is also suspect. Effective crime policy from a neurolaw perspective is not wedded to folk psychological notions of wrongfulness or even diseased minds.[222] Instead, the limits are the reach of what the behavioral sciences deems its expertise and what legislators consider amenable to social welfare policy. From this position, neurolaw's cadence of therapeutic corrections will flow freely as the idea of punishment is relegated to the province of history. Desert, and

222. *See* Bruce J Winick, *The Jurisprudence of Therapeutic Jurisprudence*, 3 PSYCHOL., PUB. POL'Y & L. 184, 193 (1997) ("Even people in the criminal justice system who may not have mental health problems per se may be thought of as having problems within the purview of therapeutic jurisprudence.").

the human intuitions from which it flows, will finally be excised from the judgments of wrongfulness.

In place of those judgments will be the measures of science. Those gauged dimensions of lawful capacity and illicit propensities will be the underpinnings of this new criminal justice system. Rather than anchored by entrenched social norms which consider some behaviors altruistic and others barbaric, it will project human experience as a random accident; the fruits of human labor as happenstance occurrences. The method of neurolaw, quite simply, understands the mind so elementarily as to see right through it entirely and instead view it solely as mechanical circuitry devoid of any of its human meaning.

What can it claim of value to the law then? It asserts that actions are blameless because of an explanation which is no explanation at all: only a description which holds motives as the exclusive domain of propagating genes and their protein offspring.[223] To say that behavior is biological is besides the point—few would seriously contend otherwise. But what the law really cares about—and has always cared about—are the reasons that people attribute to their intentions and behavior. The language of law is a mirror of humanity which reflects a rudimentary understanding that people behave because they have reasons and those reasons manifestly retain value to most people.[224]

IV. CONCLUSION

Foundational to our criminal law as commonly understood is the premise that the state cannot impose punishment where it cannot impose blame. This quality of the criminal code is the source for its unique capacity to condemn and sanction severely by depriving liberty. The act of punishment not only decrees social condemnation but incapacitates because it assumes that those who acted wrongly will choose to do so again unless they bear the full weight of shame society assigns to criminal conduct. By necessity, behavior judged criminal is thusly considered by its nature to be willful in a folk psychological

223. *See* Dawkins, *supra* note 47.

224. *Cf.* Rollin M. Perkins, *A Rationale for Mens Rea*, 52 HARV. L. REV. 905, 905 (1939) ("Deeply ingrained in human nature is the tendency to distinguish intended results from accidental happenings.").

sense.

That moral sense shared across cultures and generations also understands that people are rooted in their biological nature but not defined by it. Few routinely say that their abilities and desires are solely attributable to their constituent genetic code. Indeed, civilized society has long rejected the idea that personhood should be narrowly circumscribed by one's biology. People reject that idea not only because of its ugly past, but because most people intuitively know that they are constituted by their biology but are not identical to it.[225] There is something more to being a person than biology can tell us.

Just as well, people universally desire explanations for behavior. Most explanations leave us unsatisfied. It is a uniquely human trait to ponder endlessly why people act and behave as they do. The allure of cognitive neuroscience is its promise to make the object of that inquiry transparent.[226] But in so doing, it must remove the attributes which make it inimitably human. There is no greed or malice or goodwill, only biophysical flotsam and jetsam and the discharge of neurons.[227] And in the process, that explanation destroys the social intuitions of behavior which give the law its meaning. Blaming the brain for unlawful behavior is not blame at all. Rather, it submits justice to the efficient calculation of science itself and binds it to that pursuit exclusively. A resolute belief in the value of common intuitions is necessary to the very idea of justice which is not tyranny or obedience which is not slavery.

---

225. *See*, *e.g.,* Walter Glannon, *Our Brains Are Not Us*, 23 BIOETHICS 321, 321 (2009) ("[T]he brain is necessary but not sufficient to account for all the physiological and psychological properties that make each of us a unique person.").

226. And when matters are transparent, there is nothing left to measure. *See generally* LEONARD COHEN, THE FUTURE (Sony/ATV Songs 1992).

227. *See generally* Stephen J. Morse, *Rationality and Responsibility*, 74 S. CAL. L REV. 251, 252, 253 (2000) (discussing the limits of scientific explanations for behavior under a legal lens).