PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE

*Michael S. Pardo**
*Dennis Patterson***

According to a wide variety of scholars, scientists, and policymakers, neuroscience promises to transform law. Many neurolegalists—those championing the power of neuroscience for law—proceed from problematic premises regarding the relationship of mind to brain. In this Article, Professors Pardo and Patterson make the case that neurolegalists' accounts of the nature of mind are implausible and that their conclusions are overblown. Thus, neurolegalists' claims of the power of neuroscience for law cannot be sustained. The authors discuss a wide array of examples, including lie detection, criminal law doctrine, economic decision making, moral decision making, and jurisprudence.

TABLE OF CONTENTS

Introduction 1212
I. The Mind as Brain 1218
II. The Conceptual and the Empirical 1220
  A. Criterial and Inductive Evidence 1222
  B. The Mereological Fallacy 1225
III. The Mereological Fallacy in Law and Neuroscience 1227
  A. Lie Detection 1227
  B. Criminal Law 1231
  C. Neuroeconomics 1235
  D. Moral Decision Making 1239
  E. Jurisprudence 1242

* Associate Professor, University of Alabama School of Law. For helpful comments on drafts of this Article, we wish to thank Ron Allen, Bebhinn Donnelly, Daniel Goldberg, Peter Hacker, Andrew Halpin, Adam Kolber, and Hans Oberdiek. We also wish to thank those who provided helpful comments in response to presentations at the Stanford Law School Neuroscience Workshop, the Association of American Law Schools' Mid-Year Meeting on Evidence, and the Northwestern Legal Philosophy Club. Finally, we wish to thank the *University of Illinois Law Review* for their excellent editorial work.

** Professor of Law, European University Institute, Florence; Board of Governors Professor of Law and Philosophy, Rutgers University (Camden); Professor of Jurisprudence and International Trade, Swansea University, UK.

IV. The "Theory" of Folk Psychology 1245
V. Conceptions of Mind and the Role of Neuroscience in Law..... 1248

INTRODUCTION

Scholars, scientists, and journalists alike all proclaim the great promise of neuroscience for law.[1] Research into the neural workings of the human brain—with the aid of sophisticated brain-imaging techniques such as functional magnetic resonance imaging (fMRI)[2]—will, some predict, "probably completely change . . . nearly every area of law."[3] Some believe that, in time, neuroscience will "dominate the entire legal system."[4] The depth and breadth of such claims is without precedent.[5]

Current proposals for the confluence of law and neuroscience have focused widely from specific doctrinal issues and areas[6] to general evidentiary issues[7] to philosophical issues involving justice, morality, free-

1. The most obvious metric for measuring the growing influence of legal interest in neuroscience is financial. With funding from the MacArthur Foundation, a group of scientists, philosophers, and lawyers have been assembled to study the implications of neuroscience for law. For a description of the project, see Michael S. Gazzaniga, *The Law and Neuroscience*, 60 NEURON 412 (2008). For a general overview of the current and future legal issues to which neuroscience may be put, see Jeffrey Rosen, *The Brain on the Stand*, N.Y. TIMES, Mar. 11, 2007, § 6 (Magazine), at 49 (quoting several scientists and legal scholars); see also Robert Lee Hotz, *The Brain, Your Honor, Will Take the Witness Stand*, WALL ST. J., Jan. 16, 2009, at A7 (discussing law and neuroscience).

2. The images of the brain generated by fMRI are constructed by measuring differences in the magnetic properties in blood as it flows to different parts of the brain. *See* William G. Gibson, Les Farnell & Max R. Bennett, *A Computational Model Relating Changes in Cerebral Blood Volume to Synaptic Activity in Neurons*, 70 NEUROCOMPUTING 1674, 1674 (2007). The tests, therefore, do not measure neural activity directly; rather, blood flow is used as an indirect measure of such activity. *See id.*; *see also* Adina L. Roskies, *Neuroimaging and Inferential Distance*, 1 NEUROETHICS 19, 23–29 (2008) (discussing epistemic problems that arise when fMRI images are mistakenly analogized to photographs). Functional MRI is the latest and most sophisticated technique for examining the workings of the brain. Other techniques include electro-encephalography (EEG), positron emission tomography (PET), and single photon emission computed tomography (SPECT). For an accessible general overview of fMRI and these other techniques, see Henry T. Greely & Judy Illes, *Neuroscience-Based Lie Detection: The Urgent Need for Regulation*, 33 AM. J.L. & MED. 377, 378–84 (2007).

3. Terrence Chorvat & Kevin McCabe, *The Brain and the Law*, *in* LAW & THE BRAIN 113, 128 (Semir Zeki & Oliver Goodenough eds., 2006). It should be noted that *Law & the Brain* was originally published as an issue of the *Philosophical Transactions of the Royal Society B: Biological Sciences* (Volume 359), but has been "materially changed and updated." *See* LAW & THE BRAIN, *supra*, at viii.

4. MICHAEL S. GAZZANIGA, THE ETHICAL BRAIN 88 (2005).

5. Perhaps only the more bombastic claims made on behalf of law and economics come close.

6. *See, e.g.*, Eyal Aharoni et al., *Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience*, 1124 ANNALS N.Y. ACAD. SCI. 145 (2008) (discussing insanity and diminished capacity in criminal law); Joshua W. Buckholtz et al., *The Neural Correlates of Third-Party Punishment*, 60 NEURON 930 (2008) (discussing third-party assessment of responsibility and determination of appropriate punishment); Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts*, 87 MINN. L. REV. 269 (2002) (discussing the actus reus requirement in criminal law); Erin Ann O'Hara, *How Neuroscience Might Advance the Law*, *in* LAW & THE BRAIN, *supra* note 3, at 21, 27–30 (discussing the mens rea requirement in criminal law); Jeffrey Evans Stake, *The Property 'Instinct,' in* LAW & THE BRAIN, *supra* note 3, at 185 (discussing genetic predispositions to acquire and retain property).

7. *See, e.g.*, Jonathan A. Fugelsang & Kevin N. Dunbar, *A Cognitive Neuroscience Framework for Understanding Causal Reasoning and the Law*, *in* LAW & THE BRAIN, *supra* note 3, at 157; Charles N.W. Keckler, *Cross-Examining the Brain: A Legal Analysis of Neural Imaging for Credibility Im-*

dom, rationality, and jurisprudence.[8] Indeed, one would be hard pressed to think of a single legal issue not potentially affected by the claims made on behalf of neuroscience.

At its most general level, law regulates human behavior. Human action *tout court* is just one of many issues within the scope of the neurolaw literature.[9] Moreover, legal judgments are made by human decision makers, who act and decide based on reasons, which is just more behavior to be reduced within a neurolaw framework.[10] Given the strong claims made on behalf of neuroscience, coupled with the growing enthusiasm for the enterprise,[11] careful scrutiny is warranted. Scholarly attention is now being directed at several important issues raised by neuroscience's implications for law.[12] One set of issues concerns the research itself, the knowledge it generates, and its empirical limitations. A second set of issues revolves around the question of how to integrate our current neuroscientific knowledge with extant legal issues and doctrine. A third focus is on the ethical implications of the law's use of neuroscience in the ways contemplated in the second set. The example of neuroscience-based lie detection helps to illustrate the three categories. Some have

peachment, 57 HASTINGS L.J. 509, 537–53 (2006); Adam J. Kolber, *Pain Detection and the Privacy of Subjective Experience*, 33 AM. J.L. & MED. 433, 441–49 (2007); F. Andrew Kozel et al., *Detecting Deception Using Functional Magnetic Resonance Imaging*, 58 BIOLOGICAL PSYCHIATRY 605, 611–12 (2005).

8. *See, e.g.*, Terrence Chorvat & Kevin McCabe, *Neuroeconomics and Rationality*, 80 CHI.-KENT L. REV. 1235, 1248–54 (2005); Oliver R. Goodenough, *Mapping Cortical Areas Associated with Legal Reasoning and Moral Intuition*, 41 JURIMETRICS J. 429, 439–42 (2001) [hereinafter Goodenough, *Mapping Cortical Areas*]; Oliver R. Goodenough & Kristin Prehn, *A Neuroscientific Approach to Normative Judgment in Law and Justice*, *in* LAW & THE BRAIN, *supra* note 3, at 77, 90–95; Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, *in* LAW & THE BRAIN, *supra* note 3, at 207; Robert A. Hinde, *Law and the Sources of Morality*, *in* LAW & THE BRAIN, *supra* note 3, at 37.

9. *See* Greene & Cohen, *supra* note 8, at 208.

10. *See* Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 437–40.

11. Law's enthusiasm for neuroscience has followed an explosion of interest in the field of neuroscience itself. In a 2007 article, Carter Snead reports that "over the past five years, an average of one thousand peer-reviewed, neuroimaging-based scholarly articles have been published each month." O. Carter Snead, *Neuroimaging and the "Complexity" of Capital Punishment*, 82 N.Y.U. L. REV. 1265, 1273 (2007) (emphasis omitted). Some of the enthusiasm within the law, however, may have less to do with the epistemic support neuroscience provides than with the simplicity of the reductive explanations it can help to generate, *see* Deena Skolnick Weisberg et al., *The Seductive Allure of Neuroscience Explanations*, 20 J. COGNITIVE NEUROSCIENCE 470, 472–77 (2008) (describing a study in which irrelevant neuroscience information appeared to cause subjects to give credence to otherwise bad explanations), particularly when accompanied by colorful images of the brain, *see* David P. McCabe & Alan D. Castel, *Seeing Is Believing: The Effect of Brain Images on Judgments of Scientific Reasoning*, 107 COGNITION 343, 349–51 (2008) (describing a study in which information accompanied by brain images was found more persuasive than the same information presented without the images); *see also* Jessica R. Gurley & David K. Marcus, *The Effects of Neuroimaging and Brain Injury on Insanity Defenses*, 26 BEHAV. SCI. & L. 85, 92–94 (2008) (describing a study in which the presence of brain-image evidence significantly increased the percentage of mock jurors accepting an insanity defense).

12. The MacArthur Foundation's "Law & Neuroscience Project," *see supra* note 1, maintains a website that provides a useful overview of this research. The Law & Neuroscience Project, http://www.lawandneuroscienceproject.org (last visited May 14, 2010); *see also* WILLIAM R. UTTAL, NEUROSCIENCE IN THE COURTROOM: WHAT EVERY LAWYER SHOULD KNOW ABOUT THE MIND AND THE BRAIN xiv–xvi (2009).

scrutinized the studies purporting to detect deception through fMRI by pointing out both their empirical limitations and their lack of fit with real-world litigation situations.[13] Others have discussed (at a doctrinal level) the various evidentiary issues and procedural issues implicated by the use of neuroscience as legal evidence.[14] Finally, others have explored the ethics of the law's potential use of this evidence with regard to issues of privacy, dignity, and autonomy.[15]

Notwithstanding the broad and varied inquiries into the relation of law to neuroscience, one particular aspect of the discussion has been missing. This is a type of scrutiny that we contend is critically important. Indeed, it is fundamental. This aspect focuses on the conceptual assumptions underlying the various law-and-neuroscience claims. Examining these assumptions is of fundamental importance because they affect virtually every inference drawn from neuroscience research and every conclusion regarding its implications for law.

We examine these assumptions in this Article. We find that many of the claims made in the neurolaw literature presuppose an implausible conception of mind in general and of the relation of mind to the brain in particular.[16] Moreover, the strength of the claims made on behalf of neuroscience depends necessarily on these implausible conceptions. Once the problematic nature of these conceptions is exposed, the plausibility of the claims that rest upon them rapidly diminishes. Our arguments apply generally and broadly, but to illustrate them we focus in detail on a variety of specific issues. For the sake of length, we have chosen examples in five different areas that well exemplify the general issues. They appear to us to be among the most powerful and clearest expressions of the bold claims made in the literature on behalf of neuroscience for law.[17] These areas are (1) lie detection,[18] (2) voluntary action and mental states

13. *See, e.g.*, Greely & Illes, *supra* note 2, at 402–05; George T. Monteleone et al., *Detection of Deception Using fMRI: Better than Chance, but Well Below Perfection*, 4 SOC. NEUROSCIENCE 528, 533–37 (2009); Walter Sinnott-Armstrong et al., *Brain Images as Legal Evidence*, 5 EPISTEME 359, 359–60, 366–67, 370 (2008). On empirical limitations more generally, see Edward Vul et al., *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition*, 4 PERSP. ON PSYCHOL. SCI. 274, 280–85 (2009) (article formerly known as *Voodoo Correlations in Social Neuroscience*, *see* Vul et al., *supra*, at 274 n.1).

14. *See, e.g.*, Jane Campbell Moriarty, *Flickering Admissibility: Neuroimaging Evidence in the U.S. Courts*, 26 BEHAV. SCI. & L. 29, 30, 47–49 (2008); Michael S. Pardo, *Neuroscience Evidence, Legal Culture, and Criminal Procedure*, 33 AM. J. CRIM. L. 301 (2006).

15. *See, e.g.*, Paul Root Wolpe, Kenneth R. Foster & Daniel D. Langleben, *Emerging Neurotechnologies for Lie-Detection: Promises and Perils*, AM. J. BIOETHICS, Mar.–Apr. 2005, at 39, 45–46.

16. *See infra* Part V. Our critique, of course, is not meant to be directed against everyone writing within neurolaw. We recognize the great deal of diversity and nuance among the various arguments made, and positions taken, within neurolaw, including many other critical voices. What unites the targets of our critique are bold claims about the power of neuroscience to solve or transform law, legal issues, or legal theory, while relying on problematic conceptual presuppositions to support those claims.

17. We also note several additional examples along the way.

18. *See* Lawrence A. Farwell & Sharon S. Smith, *Using Brain MERMER Testing to Detect Knowledge Despite Efforts to Conceal*, 46 J. FORENSIC SCI. 135, 141–43 (2001); G. Ganis et al., *Neural Correlates of Different Types of Deception: An fMRI Investigation*, 13 CEREBRAL CORTEX 830 (2003);

in criminal law,[19] (3) economic decision making,[20] (4) moral decision making,[21] and (5) jurisprudential questions regarding the nature of law and legal reasoning.[22]

Before outlining our arguments, however, we wish to dispel a potential confusion at the outset. Our talk of "implausible" or "problematic" concepts of "mind" may suggest to readers that we are setting up a classic dualist versus materialist discussion, with the neuroscience proponents falling on the materialist side.[23] To be clear, nothing could be further from the truth. Indeed, as we will discuss, this putative dichotomy is a principal source of the problem. Cartesian dualism—with its picture of mind as a mysterious immaterial substance, independent of but in causal relation with the body[24]—is typically set up as the foil in neuroscience discussions. For example, in introducing the journal *Neuroethics*, Neil Levy writes that "Cartesian (substance) dualism is no longer taken seriously; the relation between the brain and the mind is too intimate for it to be at all plausible. . . . [N]euroscientific discoveries promise . . . to reveal the structure and functioning of our minds and, therefore, of our souls."[25] Likewise, in discussing the implications of neuroscience for jurisprudence, Oliver Goodenough writes that the "Cartesian model . . .

Keckler, *supra* note 7, at 537–40; Kozel et al., *supra* note 7; F. Andrew Kozel et al., *A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men*, 16 J. NEUROPSYCHIATRY & CLINICAL NEUROSCIENCES 295, 302–04 (2004); D.D. Langleben et al., *Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study*, 15 NEUROIMAGE 727, 730–31 (2002); Tatia M.C. Lee et al., *Lie Detection by Functional Magnetic Resonance Imaging*, 15 HUM. BRAIN MAPPING 157, 161–63 (2002); Sean A. Spence et al., *A Cognitive Neurobiological Account of Deception: Evidence from Functional Neuroimaging*, *in* LAW & THE BRAIN, *supra* note 3, at 169, 174–79.

19. Denno, *supra* note 6, at 320–37; O'Hara, *supra* note 6, at 28–30.

20. Chorvat & McCabe, *supra* note 3, at 120–24; Chorvat & McCabe, *supra* note 8, at 1248–54; Alan G. Sanfey et al., *The Neural Basis of Economic Decision-Making in the Ultimatum Game*, 300 SCIENCE 1755 (2003) [hereinafter Sanfey at al., *Ultimatum*]; Alan G. Sanfey et al., *Neuroeconomics: Cross-Currents in Research on Decision-Making*, 10 TRENDS COGNITIVE SCI. 108, 111 (2006) [hereinafter Sanfey et al., *Neuroeconomics*]; *see also* Morris B. Hoffman, *The Neuroeconomic Path of the Law*, *in* LAW & THE BRAIN, *supra* note 3, at 3, 11–14; Jedediah Purdy, *The Promise (and Limits) of Neuroeconomics*, 58 ALA. L. REV. 1, 13–15, 39–40 (2006); Paul J. Zak, *Neuroeconomics*, *in* LAW & THE BRAIN, *supra* note 3, at 133, 135–36.

21. Chorvat & McCabe, *supra* note 3, at 118–20; Joshua Greene, *From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?*, 4 NATURE REVS. NEUROSCIENCE 847 (2003); Joshua Greene & Jonathan Haidt, *How (and Where) Does Moral Judgment Work?*, 6 TRENDS COGNITIVE SCI. 517 (2002); Joshua D. Greene et al., *Cognitive Load Selectively Interferes with Utilitarian Moral Judgment*, 107 COGNITION 1144 (2008); Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCIENCE 2105 (2001) [hereinafter Greene et al., *An fMRI Investigation*]; Joshua D. Greene et al., *The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 44 NEURON 389 (2004) [hereinafter Greene et al., *Neural Bases*].

22. Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 339–41.

23. According to prominent neuroscientist Michael Gazzaniga, "'98 or 99 percent'" of cognitive neuroscientists subscribe to the reduction of mind to the brain in their attempts to explain mental phenomena. Snead, *supra* note 11, at 1279 (quoting Michael Gazzaniga, *quoted in* Richard Monastersky, *Religion on the Brain*, CHRON. HIGHER EDUC. (Wash., D.C.), May 26, 2006, at A15).

24. For an overview of this position, see Howard Robinson, *Dualism*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2009), http://plato.stanford.edu/entries/dualism/.

25. Neil Levy, *Introducing* Neuroethics, 1 NEUROETHICS 1, 2 (2008) (emphasis omitted).

supposes a separation of mind from the brain," whereas models of mind for "a nondualist like myself" are "what the brain does for a living."[26] The dichotomy between dualism and mind-as-brain is a false one. Moreover, as we will discuss, materialists like Goodenough are *too* Cartesian—he, like many neuroscientists and neurolaw scholars have (ironically) kept the problematic Cartesian structure in place by simply replacing the Cartesian soul with the brain.[27] We offer a more plausible alternative.

Rather than arguing about where the mind is located (e.g., in the brain or elsewhere) we need to step back and contemplate whether this is the right question to ask. First, notice that the question of the mind's location presupposes that the mind is a kind of "thing" or "substance" that is located "somewhere" (e.g., in the body). Why must this be so? Our answer is that it need not be, and is not. An alternative conception of mind—the one that we contend is more plausible—is as a diverse array of abilities exercised by a person.[28] These abilities implicate a wide range of psychological categories including sensations, perceptions, cognition (i.e., knowledge, memory), cogitation (i.e., beliefs, thought, imagination, mental imagery), affectations (i.e., emotions, moods, appetites), and volition (i.e., intentions, voluntary action).[29]

To be clear, we do not deny that a properly working brain is necessary for a person to engage in the diverse array of abilities that we collectively identify as mental life. Again, nothing could be further from the truth. But, while certain neural activity is *necessary* for a human being to exercise these abilities, neural activity alone is not *sufficient*. The criteria for the successful exercise of an ability are not a matter of what is or is not in the mind or brain. These criteria—which are normative in nature—are the basis for our attribution of mental states to one another.[30]

26. Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 431–32.

27. *See* M.R. BENNETT & P.M.S. HACKER, PHILOSOPHICAL FOUNDATIONS OF NEUROSCIENCE 233–35 (2003) (tracing the *explicit* Cartesian structure of mind in early neuroscience through its transformation to an *implicit* Cartesian structure around the late nineteenth and early twentieth centuries).

28. *See id.* at 62–63 ("The mind, as we have already intimated, is not a substance of any kind. . . . We say of a creature (primarily of a human being) that it *has a mind* if it has a certain range of active and passive powers of intellect and will—in particular, conceptual powers of a language-user that make self-awareness and self-reflection possible.").

29. We do not mean to imply that all of the categories we enumerate are to be understood as being on a par with one another or classifiable under one scheme. For example, what is distinctive about the abilities of creatures with a mind is that they can and do act for reasons. Thus, abilities of intellect and will are to be distinguished from those of sensation and perception. Each category requires its own detailed analysis. In this Article, we provide a detailed account of a few of these in the course of making a larger point about the complexity of mental life in the face of the reductionist impulses of the neurolaw devotees.

30. *See* DONALD DAVIDSON, *Three Varieties of Knowledge*, *in* A.J. AYER: MEMORIAL ESSAYS 153 (A. Phillips Griffiths ed., 1991), *reprinted in* SUBJECTIVE, INTERSUBJECTIVE, OBJECTIVE 205, 207 (2001) ("No doubt it is true that it is part of the concept of a mental state or event that behavior is evidence for it."). To avoid another potential confusion, note that we are not behaviorists either. While psychological capacities are manifested in behavior (and thus the behavior provides evidence of them) we are not suggesting, as a behaviorist would, that the capacities are identical with or can be reduced to behavior. Unlike behaviorists, we acknowledge that psychological events may sometimes take

To outline briefly one of the examples that we will explore below, consider what it means to "have knowledge." We believe that "knowing" is not (just) having a brain in a particular physical state. Rather, it is having the ability to do certain things (e.g., to answer questions, correct mistakes, act correctly on the basis of information, and so on). Thus, if behavior, and not brain states, constitutes the criteria for "knowledge," then it will make no sense[31] to say that knowledge is "located" in the brain. The same is true for other psychological predicates—and for the mind itself. So, to the question "where is the mind located—the Cartesian soul or the brain?"—we think the best answer is "neither." The question of the mind's location makes no sense. Humans have minds, but minds are not *in* them.[32]

We recognize that our claims may initially strike those operating within the dualist versus mind-as-brain dichotomy, or those relying on common-sense notions about the mind, as fanciful (or worse). Thus, to undermine the entrenched but problematic assumptions underlying the neurolaw claims, we proceed deliberately and carefully. After first presenting an example of the problematic conception of mind in the neurolaw literature,[33] we begin our argument by introducing an important methodological distinction between conceptual and empirical questions. In the context of neuroscience research, empirical claims are those that are amenable to confirmation or falsification on the basis of experiments or data.[34] By contrast, conceptual questions concern the logical relations between concepts.[35] We explain why the question of what the mind *is* and what the various psychological categories under discussion *are* (e.g., knowledge, intention, rationality), are conceptual rather than empirical questions.

Given that these are conceptual issues, we next discuss two related conceptual issues that are crucial for evaluating the neurolaw claims and arguments: the distinction between criterial and inductive evidence, and the mereological fallacy (a fallacy relating to the logic of part/whole relations).[36] The first concerns the inferences that may be drawn from a body of evidence (neuroscience research) regarding various capacities and their exercise; the second considers whether it makes sense to ascribe psychological predicates to the brain, rather than to the person as a

place in the absence of behavior and that behavior may take place in the absence of psychological events.

31. We discuss the "sense" of claims in more detail in Part II. *See infra* notes 54–57 and accompanying text.

32. Unless we wish to speak metaphorically, as in when we ask whether he "has it in him" to win the game.

33. *See infra* Part I; *see also* Greene & Cohen, *supra* note 8.

34. *See infra* Part II.

35. *See infra* Part II.

36. *See infra* Part II.A–B.

whole.[37] With these important methodological and substantive issues in place, we then turn to several issues within the neurolaw literature. We illustrate how problematic assumptions made in the literature—confusions regarding criterial versus inductive evidence and instances of the mereological fallacy—permeate and ultimately render implausible many of the claims made on behalf of neuroscience for law.

We then turn to a further deep challenge raised within the neurolaw literature: the eliminative-materialist challenge to so-called "folk psychology."[38] This challenge turns on the claim that, as an empirical matter, our ordinary psychological categories (e.g., beliefs, desires, intentions)—what is dubbed the "model" or "theory" of "folk psychology"—are causally inefficacious and thus illusory. Because exercising these psychological capacities does no causal work, so the argument goes, we ought to stop using our ordinary psychological vocabulary and adopt the purely materialist assumptions of the hardcore mind-as-brain crowd. As we discuss, this eliminative-materialist challenge misses its mark.

We conclude by returning to the concept of mind, offering an alternative and the role for neuroscience in law that it implies.

I. THE MIND AS BRAIN

The typical conception of the mind in the neuroscience literature is simply that the mind is the brain.[39] The phenomena that we associate with mental life (e.g., thinking, believing, knowing) are equated with the workings of the brain or particular brain structures. For this reason, the mind is "reduced" to the brain in the sense that the operations of the mind are held to be fully explicable in terms of the operations of the brain.[40] Once this reduction takes place, there is nothing about the mind left to explain or understand.

Although many argue for the reduction of mind to brain, Joshua Greene and Jonathan Cohen have presented one of the clearest and most sophisticated arguments on behalf of reductionism and its implica-

37. These include the wide range of capacities noted above regarding sensation, perception, cognition, cogitation, affectation, and volition.

38. *See* Greene & Cohen, *supra* note 8, at 218–21.

39. *See* Rosen, *supra* note 1, at 52 (quoting Joshua Greene: "To a neuroscientist, you are your brain . . . .").

40. Reduction is a general philosophical strategy in scientific theorizing. For an overview, see Ingo Brigandt & Alan Love, *Reductionism in Biology*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 24, http://plato.stanford.edu/entries/reduction-biology/. *See also* David J. Chalmers & Frank Jackson, *Conceptual Analysis and Reductive Explanation*, 110 PHIL. REV. 315, 350–56 (2001). A more radical form of reductionism regarding psychological concepts, "eliminative materialism," seeks not only to explain them in terms of the brain, but rather to show they do not exist. *See* Paul M. Churchland, *Eliminative Materialism and the Propositional Attitudes*, 78 J. PHIL. 67 (1981); William Ramsey, *Eliminative Materialism*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 24, http://plato.stanford.edu/entries/materialism-eliminative/. We turn to this more radical position in Part IV *infra*.

tions for law.[41] According to Greene and Cohen, the most important impact of neuroscience on law lies in the future.[42] What is now emerging—and will be confirmed in the future—is our "understanding of the mind as brain."[43] The reason this new understanding of mind is so important, they maintain, is that it will transform our understanding of the relationship between mind and law.[44] By identifying the neural bases of behavior, neuroscience will transform our current understanding (they call it "folk psychology"[45]) of free will, responsibility, and human action.

Greene and Cohen use the topic of criminal responsibility to illustrate their claim that neuroscience will alter our intuitions about what it means to be "responsible for" an act or crime.[46] In the specific context of diminished capacity, "what many people really want to know is: was it really *him*?"[47] In other words, can an accused's conduct be attributed to his "circumstances" (e.g., upbringing or genes) or was it "really *him*?" And who is "him"? The answer to this question reveals Greene and Cohen's stance on the mind. They write:

> Was it *him*, or was it his *circumstances*? Was it *him*, or was it his *brain*? But what most people do not understand, despite the fact that naturalistic philosophers and scientists have been saying it for centuries, is that there is no 'him' independent of these other things. (Or, to be a bit more accommodating to the supernaturally inclined, there is no 'him' independent of these things that shows any sign of affecting anything in the physical world, including his behaviour.)[48]

For Greene and Cohen, there are just two accounts of "mind": dualist and materialist. For most people (and their intuitions), dualism—"the view that mind and brain are separate, interacting, entities"[49]—makes sense. We are responsible for our actions (and accountable for our crimes) because we have *chosen* to do what we have done. By contrast, a materialist explanation of behavior sees the role of "mind" as just one physical piece of the explanatory/causal narrative.[50] When a materialist answers the question "Was it *him*, or was it his *brain*?," the answer is "they are one and the same." For the materialist, you *are* your brain.[51]

41. *See* Greene & Cohen, *supra* note 8.
42. *See id.* at 207–08.
43. *Id.* at 207.
44. *Id.* at 213 ("The legitimacy of the law itself depends on its adequately reflecting the moral intuitions and commitments of society. If neuroscience can change those intuitions, then neuroscience can change the law.").
45. *Id.* at 220–21. We return to a discussion of "folk psychology" in Part IV *infra*.
46. Greene & Cohen, *supra* note 8, at 213–17.
47. *Id.* at 213.
48. *Id.* at 213–14.
49. *Id.* at 214.
50. Greene and Cohen posit that with a scanner of sufficient resolution you can even watch the neurons in your brain decide whether you will have soup or salad for lunch. *Id.* at 218. "You find the tipping-point moment at which the blue neurons in your prefrontal cortex out-fire the red neurons, seizing control of your pre-motor cortex and causing you to say, 'I will have the salad, please.'" *Id.*
51. Greene and Cohen spell out their view with another clear example:

Are dualism and materialism the only two conceptions of mind available? We think not. But before we explain the alternative to this false choice, we need to explain what is wrong with the materialist conception of mind. To understand what is wrong with the materialist *conceptual* assumptions regarding the mind, we must first turn to the important distinction between conceptual and empirical issues.

II. THE CONCEPTUAL AND THE EMPIRICAL

The important issue of the relationship between conceptual and empirical claims has, unfortunately, received little direct attention in the current debate over the present and future role of neuroscience in law. Empirical neuroscientific claims, and the inferences and implications for law drawn from them, depend on conceptual presuppositions regarding the mind. As we see it, many of the proponents of an increased role for neuroscience in law rest their case on a controversial and ultimately untenable account of the nature of mind. Although we recognize the need for greater emphasis on and interrogation of the empirical claims regarding neuroscience applications in law, we believe that the fundamental conceptual issues regarding the mind are of equal, if not greater, importance.

Devoted as they are to understanding the physiology of the brain, neuroscientists are principally interested in physical processes.[52] Of greatest interest to neuroscientists are questions regarding neural structures, the functioning of the brain, and the physiological bases for a wide range of human abilities, including consciousness, memory, vision, and emotion. Scientific explanations, including those of neuroscience, are framed in a language of explanation most readily identified as "empirical." Grounded in theories and hypotheses, scientific claims are tested by means of experiment. Experimental confirmation or disconfirmation of hypotheses forms the basis of the scientific method.

> When a composer conceives a symphony, the only way he or she can present it to the public is through an orchestra . . . If the performance is poor, the fault could lie with the composer's conception, or the orchestra, or both . . . Will is expressed by the brain. Violence can be the result of volition only, but if a brain is damaged, brain failure must be at least partly to blame.
>
> To our untutored intuitions, this is a perfectly sensible analogy, but it is ultimately grounded in a kind of dualism that is scientifically untenable. It is not as if there is you, the composer, and then your brain, the orchestra. You are your brain, and your brain is the composer and the orchestra all rolled together. There is no little man, no 'homunculus', in the brain that is the real you behind the mass of neuronal instrumentation. Scientifically minded philosophers have been saying this ad nauseum . . . .

Id. at 214 (alternation in original) (citation omitted) (emphasis omitted) (quoting JONATHAN H. PINCUS, BASE INSTINCTS: WHAT MAKES KILLERS KILL? 128 (2001)).

52. *See* M.R. BENNETT AND P.M.S. HACKER, HISTORY OF COGNITIVE NEUROSCIENCE 1 (2008) ("Neuroscience is concerned with understanding the workings of the nervous system . . . .").

Empirical and conceptual questions are distinct. We would go so far as to say that they are logically distinct.[53] In addition to their distinct characters, the conceptual relates to the empirical in a certain way: the very success of empirical inquiry depends upon well-founded conceptual claims. An experiment grounded in dubious conceptual claims can prove nothing.[54]

Conceptual questions concern the logical relations between concepts. Concepts like *mind*, *consciousness*, *knowledge*, and *memory* are exemplary instances of the sorts of concepts implicated in neuroscience discussions. To be well-founded, and thus to ground successful empirical claims, conceptual claims have to make sense.[55] But what does it mean to say that conceptual claims have to make "sense"?[56] The concept of sense is bound up with the forms of expression for the use of words in a language. Therefore, to say that a particular claim lacks sense (literally, is nonsense) is not to say the claim is improbable, false, or stupid; it is to say that it fails to express something meaningful. Often mistakes in use can generate "nonsensical" claims—for example, what is meant by one's (non-metaphorical) claim that a moral obligation weighs the same as an elephant? Sometimes mistakes in usage can take the form of simple grammatical errors—compare "he has almost finished his breakfast" with "he has not already finished his breakfast." More importantly, however, they sometimes ramify in more problematic and significant ways. One

53. By "logically distinct" we mean that one is not reducible to the other or explicable in its terms. The relationship between empirical and conceptual claims is itself a matter of philosophical controversy. The latest philosophical trend—"experimental philosophy"—blends empirical methods and conceptual investigations. A representative collection of papers may be found in EXPERIMENTAL PHILOSOPHY (Joshua Knobe & Shaun Nichols eds., 2008).

54. For example, imagine an experiment purporting to "prove" that moral obligations weigh more than elephants. We do not mean to suggest that an experiment based on faulty conceptual presuppositions cannot sometimes produce fruitful results that do not depend on the faulty conceptual presuppositions.

55. BENNETT & HACKER, *supra* note 27, at 148 ("A prerequisite for fruitful and illuminating empirical research on the neural foundations of our psychological capacities is clarity concerning the concepts involved.").

56. Bennett and Hacker explain the relationship of sense to truth thus:

> Cognitive neuroscience is an experimental investigation that aims to discover empirical truths concerning the neural foundations of human faculties and the neural processes that accompany their exercise. A precondition of truth is sense. If a form of words makes no sense, then it won't express a truth. If it does not express a truth, then it can't explain anything. Philosophical investigation into the conceptual foundations of neuroscience aims to disclose and clarify conceptual truths that are presupposed by, and are conditions of the sense of, cogent descriptions of cognitive neuroscientific discoveries and theories. If conducted correctly, it will illuminate neuroscientific experiments and their description as well as the inferences that can be drawn from them. In *Philosophical Foundations of Neuroscience* we delineated the conceptual network formed by families of psychological concepts. These concepts are presupposed by cognitive neuroscientific research into the neural basis of human cognitive, cogitative, affective, and volitional powers. If the logical relations of implication, exclusion, compatibility, and presupposition that characterize the use of these concepts are not respected, invalid inferences are likely to be drawn, valid inferences are likely to be overlooked, and nonsensical combinations of words are likely to be treated as making sense.

Maxwell Bennett & Peter Hacker, *The Conceptual Presuppositions of Cognitive Neuroscience: A Reply to Critics*, *in* NEUROSCIENCE AND PHILOSOPHY: BRAIN, MIND AND LANGUAGE 127, 128 (2007) (footnote omitted).

such mistake occurs when we think of properties like "mind" as if they must be a type of substance (like the brain).[57]

This error underlies the fundamental reductionist move in the positive argument for neuroscience in law.[58] As discussed above, the reduction is the reduction of the mind to the brain. By making this move, proponents of an increased role for neuroscience set the stage for their enterprise, which is the explanation of human behavior in causal, mechanical, or non-volitional terms.[59] As we will show, the reductive impulse is driven by a conceptually problematic account of the relationship between mind and brain. Once this account is undermined, many of the aspirations of the neurolegalists diminish significantly, if not entirely. As mentioned, we expose the problematic foundations of these accounts by focusing on two related conceptual issues: the distinction between criterial and inductive evidence, and the mereological fallacy. We first explicate the conceptual issues and then turn to several legal examples, illustrating in detail the conceptual problems underlying the diverse arguments made on behalf of, and inferences drawn from, neuroscience research for law.

A. Criterial and Inductive Evidence

Suppose we were asked to look for evidence of various kinds of psychological faculties or attributes such as perception and belief. Some evidence would provide *criterial* support—that is, it would provide constitutive evidence for the faculty or attribute.[60] Another class of evidence would provide *inductive* support—that is, although not constitutive of the faculty or attribute, it might be empirically well-correlated with the faculty or attribute so that we could say with some degree of confidence that the presence of this evidence increases (or decreases) the likelihood of the phenomena with which it is correlated.[61]

Criterial evidence for the ascription of psychological predicates, such as "to perceive" or "to believe," consists in various types of behavior.[62] Behaving in certain ways is logically good evidence and, thus,

57. Bennett and Hacker trace the genealogy of this mistake in BENNETT & HACKER, *supra* note 27, at 324–28.

58. As noted above, a prominent neuroscientist claims that "98 or 99 percent" of cognitive neuroscientists subscribe to the reduction of mind to the brain in their attempts to explain mental phenomena. *See supra* note 23.

59. *See* Snead, *supra* note 11, at 1273–74; *see also supra* notes 6–8 and accompanying text.

60. For discussion of the concept of "criteria," see LUDWIG WITTGENSTEIN, THE BLUE AND BROWN BOOKS 24–25 (1958). For general explication of Wittgenstein on criteria, see JOACHIM SCHULTE, WITTGENSTEIN: AN INTRODUCTION 130–32 (William H. Brenner & John F. Holley trans., 1992).

61. *See* BENNET & HACKER, *supra* note 27, at 68–70; *see also* James Hawthorne, *Inductive Logic*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 24, http://plato.stanford.edu/entries/logic-inductive/.

62. As noted above, while we endorse the view that behavior plays a central role in the formation of criteria, we do not ascribe to the behaviorist account of human action. *See supra* note 30.

partly constitutive of these concepts. For visual perception, this includes, for example, that one's eyes track the phenomena one perceives, that one's reports match what one observed, and so on.[63] For belief, this includes, for example, that one asserts or endorses what one believes, that one acts in ways consistent with one's beliefs, that one does not believe directly contradictory propositions, and so on.[64] This behavior is not only a way to determine whether someone perceives or believes something in particular. The behavior also helps to determine (it partly *constitutes*) what it means to engage in these activities. In other words, it helps to provide the *measure* for whether someone is in fact engaged in this activity (not just a measurement in a particular instance).[65] If these forms of behavior were not possible for a creature, then it would not make sense to ascribe the predicate to it truly or falsely.[66] Note, however, that this criterial evidence is *defeasible*; people can assert propositions they do not believe, or say they perceived things they did not, and people can perceive or believe without ever describing what they perceived or asserting or acting on what they believe. The primary point is that the behavior not only provides evidence of whether someone on a given occasion is

Roughly speaking, behaviorists reduce psychological states to behavior in explaining human action. We think the better explanatory course is to show how behavior is woven together with the language of the mental in the formation of criteria for the ascription of psychological states such as belief, desire, and intention. In this regard, we follow the philosophical approach of Gilbert Ryle and Ludwig Wittgenstein. Ryle and Wittgenstein undermined the Cartesian picture of mind by attacking the fundamental assumption underlying that picture, to wit, the notion that "the mind" is an inner theater that is properly the object of research. The inner/outer dichotomy lies at the heart of Cartesianism and its manifold confusions. For a discussion of Cartesianism and Behaviorism, see Wes Sharrock & Jeff Coulter, *ToM: A Critical Commentary*, 14 THEORY & PSYCHOL. 579, 582–87 (2004).

63. Of course, they sometimes may be mistaken on particular instances or even systematically mistaken (e.g., one who is colorblind). But if the reports did not appear to have any connection with what was happening around a person, we would not say that the person was *perceiving* anything. *See* BENNETT & HACKER, *supra* note 27, at 127 ("[T]he forms of behaviour that manifest possession of a given perceptual faculty consist in relative efficiency in discrimination, recognition, discernment, pursuit of goals and exploration of the environment, and, in the case of human beings, in corresponding utterances. These kinds of behaviour in response to visibilia, for example, are logical criteria for a creature's seeing things.").

64. Again, particular instances may create exceptions, but wholesale failures would cause us to question whether the person actually held the beliefs purportedly attributed to them. This is also why assertions such as "P, but I don't believe P" (Moore's paradox) generally are considered to be contradictory. *See* Roy Sorensen, *Epistemic Paradoxes 5.3*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 24, http://plato.stanford.edu/entries/epistemic-paradoxes/#MooPro.

65. *See* Bennett & Hacker, *supra* note 56, at 130 ("[T]o characterize a sentence as expressing a conceptual truth is to single out its distinctive function as a statement of a measure, rather than of a measurement." (emphasis omitted)).

66. Or, alternatively, such uses may be intended to change the meaning of "perceiving" or "believing." We should note that there is nothing problematic about scientists, philosophers, law professors, or anyone else coining new terms or putting existing terms to new purposes. The conceptual problems we are discussing with regard to neuroscience and neurolaw claims occur because the claims purport to tell us about our extant, ordinary psychological faculties and attributes (such as believing, perceiving, and knowing)—not because the authors are coining new terms or extending existing ones.

perceiving or has a belief, it also partially determines what it means to perceive or believe.[67]

By contrast, some evidence provides only inductive support for whether one is perceiving or believing. This would be the case if there were, as an empirical matter, a correlation between some evidence and perceiving or believing. Neural activity, as demonstrated by neuroscience research, may fill this role; searching for these correlations is precisely the goal of much current research.[68] But note that this inductive correlation only works once we know what to correlate the neural activity *with*.[69] Physical states of the brain are not criterial evidence for—because they are not partly constitutive of—psychological faculties and attributes such as perception or belief.[70] To refer back to the metaphor in the above paragraph, neural activity may help to provide a *measurement*—but not the *measure*—of whether one has perceived or believes something on a particular occasion.[71]

To know whether a brain state is correlated with a particular psychological faculty or attribute, we must first have criteria for *identifying* the faculty or attribute. Physical states of the brain cannot fulfill this role. To illustrate this, consider a claim that a certain brain state, or pattern of neural activity, constitutes perceiving *X* or thinking *that P is true*, but that a person whose brain was in either of these states engaged in none of the behavior that we associate with thinking or perceiving.[72] Suppose we ask the person and she sincerely denies that she had per-

67. *See* Richard Rorty, *The Brain as Hardware, Culture as Software*, 47 INQUIRY 219, 231 (2004) ("[B]eliefs cannot be individuated in such a way as to correlate with neural states.").

68. *See, e.g.*, Maxwell Bennett, *Epilogue* to NEUROSCIENCE AND PHILOSOPHY: BRAIN, MIND, AND LANGUAGE, *supra* note 56, at 163 (discussing the neuroscience of perception).

69. An analogy may be helpful here. A bounty hunter chasing a fugitive is no doubt interested in catching the fugitive and not the fugitive's picture on the "wanted" poster. But the bounty hunter ought to pay attention to the details on the poster in order to identify the fugitive, to know whom to look for. Likewise, even though neuroscientists and legal scholars may be interested in our psychological capacities rather than our concepts for those capacities, they ought to pay attention to the details of these concepts in order to help search for and identify these capacities. This analogy is taken from FRANK JACKSON, FROM METAPHYSICS TO ETHICS: A DEFENCE OF CONCEPTUAL ANALYSIS 30–31 (2000).

70. If neural activity did provide criterial evidence, then having particular brain states would constitute exercising the ability (perceiving) or having the attribute (believing). *Cf.* BENNETT & HACKER, *supra* note 27, at 173–74 ("There are sceptical and gullible people, but no sceptical and gullible brains. We all know what it is for a person to believe or not to believe in God, to believe in the Conservative Party or in fairies, to believe a person or his story or to doubt a person's word and be sceptical about his story. But we do not know what a religious, agnostic or atheist *brain* might be. No sense has been given to such a form of words.").

71. The same proposition may serve as a *measure* in one context and a *measurement* in another context. The difference depends on whether it is being used in a normative, constitutive role or a purely descriptive one.

72. Certain neural activity may be necessary to engage in (and play a causal role in) the behavior that constitutes the ability to think or perceive, and neuroscientists may discover this relationship by examining correlations between brain states and neural activity. But, again, this would show only that such activity was a *necessary* condition, not a *sufficient* condition, for abilities such as perceiving or believing. The behavior with which the activity was correlated would still be what provided the criterial evidence.

ceived or thought anything. In this example, the claim that the particular brain states constitute thinking or perceiving would be false, based in part on the constitutive evidence to the contrary (her sincere denial).[73] Any purported inductive correlation between the particular brain states and thinking or perceiving would have to be reexamined.

B. The Mereological Fallacy

If anything unites the various problems and projects of neurolegalists, it is the belief that the mind and the brain are one. This belief has spread far beyond neurolegalists, for it is a pervasive feature of much of the current literature and research in neuroscience as well as more popular writings.[74] But does it make sense to attribute to the brain psychological attributes normally attributed to persons? Can we intelligibly say that the brain thinks, perceives, feels pain, and decides? If we cannot, what are the implications for neuroscience and law?

Our argument that neurolegalists commit the "mereological fallacy" begins with the conceptual-empirical distinction. As discussed earlier, two distinct sorts of questions permeate discussions of mental life. Empirical questions are the focus of scientific research, specifically research into the biology and physiology of brain function.[75] By contrast, conceptual questions address how the relevant concepts are articulated. At all times, the point of the philosophical enterprise is to assess the degree to which articulations regarding the brain make sense.[76]

The mereological fallacy consists in attributing an ability or function to a part that is only properly attributable to the whole of which it is a part.[77] In this case, "part" and "whole" are elements of human beings.[78] But why is it an error—indeed a "conceptual" error—to ascribe a psychological attribute to a part of a human being? Consider, once again, "knowledge." Does the claim that knowledge is located in the brain transgress the bounds of sense so that we can say that it "makes no sense" to say that "the brain stores knowledge"? Can knowledge be stored in a brain just as information is stored in books or hard drives?

73. An additional example regarding the criterial-inductive distinction involves mental imagery. The (defeasible) criterial evidence for whether one has a particular mental image is the person's say-so and how they visualize the image. Neural evidence accompanying such mental imagery may be inductively correlated with such imagery, but having the neural events is not the criteria for having the images. *See* BENNETT & HACKER, *supra* note 27, at 187–98 for a discussion of this issue. The issue of mental imagery may have legal relevance to the issue of eyewitness identification. *See* Rosen, *supra* note 1, at 50–51 (citing Professor Owen Jones on the potential relevance of neuroscience to facial recognition).

74. *See supra* note 1 and Part I.

75. *See* BENNETT & HACKER, *supra* note 52.

76. Bennett & Hacker, *supra* note 56.

77. *See* Achille Varzi, *Mereology*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 24, http://plato.stanford.edu/entries/mereology/; *see also Foreword*, 103 J. PHIL. 593 (Wolfgang Mann & Achille C. Varzi eds., 2006).

78. *See* Bennett & Hacker, *supra* note 56, at 133–34.

In their critique of the work of philosopher Daniel Dennett, neuroscientist Maxwell Bennett and philosopher Peter Hacker argue that "[i]n the sense in which a human being possesses information, the brain possesses none."[79] Imagine the schedule for the New York Philharmonic is "encoded" in your brain. Can we say of you that you know when the next Mahler symphony is to be performed by the orchestra? If the question "when is the next Mahler symphony to be performed by the New York Philharmonic?" is put to you and you utter the wrong date, we would conclude—correctly—that you did not know the answer to the question. Knowing is not being in a particular state.[80] Knowing is an ability—the ability, for example, to answer the question correctly. The measure of the truth of your answer is not found in the neural state of your brain. Whether or not you know the answer to the question is shown, among other ways, by what you sincerely say in response to the question.[81]

The upshot of this and countless other examples is that psychological attributes are essentially manifested in the behavior, reactions, and responses of the living human being in the stream of life, not in whatever concomitant neural activity of his brain is to be found. This is the key to the mereological fallacy and the undoing of the reductive impulses of neurolegalists. Behavior is something only a human being (or other animal) can engage in. Brain functions and activities are not behaviors (and persons are not their brains). Yes, it is necessary that one have a brain in order to engage in behavior.[82] But the reduction of a psychological attribute to a cortical attribute is a fallacious move from whole to part.

If the error of ascribing attributes of a whole to one of its parts is indeed a central error in the neurolegalist agenda, what are the implica-

79. *Id.* at 137 ("In the sense in which a book contains information, the brain contains none. In the sense in which a human being possesses information, the brain possesses none.").

80. As Anthony Kenny reminds us, "To contain information is to be in a certain state, while to know something is to possess a certain capacity." ANTHONY KENNY, THE LEGACY OF WITTGENSTEIN 129 (1984).

81. Or it may be manifested in other behavior, for example, by showing up on time for the symphony. Although knowledge is typically manifested in behavior, this is not to deny that someone may lose the ability to manifest their knowledge in some ways (e.g., one who knows how to play tennis but is no longer physically able) or perhaps in all ways (e.g., someone with total "locked in" syndrome). On the latter see JEAN-DOMINIQUE BAUBY, THE DIVING BELL AND THE BUTTERFLY (1997); ALVA NOË, OUT OF OUR HEADS 14–17 (2009).

82. While having a brain is, of course, necessary, it is a mistake to suppose that having particular brain states is *sufficient*. This criticism is an across-the-board assault on the explanatory power of neuro-reductionism. Raymond Tallis explains what lies at the heart of the impulse and why it fails as an explanation of behavior:

> The appeal to brain science as an explain-all has at its heart a myth that results from confusing necessary with sufficient conditions. Experimental and naturally occurring brain lesions have shown how exquisitely holes in the brain are correlated with holes in the mind. Everything, from the faintest twinge of sensation to the most elaborately constructed sense of self, requires a brain; but it does not follow from this that neural activity is a sufficient condition of human consciousness, even less that it is identical with it.

Raymond Tallis, *License My Roving Hands*, TIMES LITERARY SUPPLEMENT (London), Apr. 11, 2008, at 13.

tions? We suggest the implications are manifold. Most importantly, the neurolegalist reduction of psychological attributes to brain states must be rejected as fallacious. Thus, voluntary action, intentionality, knowledge, and decision making cannot be attributes of brains, but only of human beings. We now illustrate these implications.

III. THE MEREOLOGICAL FALLACY IN LAW AND NEUROSCIENCE

The examples we now discuss run the gamut from discrete doctrinal issues to the nature of law itself. They include voluntary human action as well as legal, moral, and economic decision making. The common theme running through the examples is, first, the identification of some type of human behavior (e.g., deception, voluntary movements, decision making) or mental faculty or attribute (e.g., knowledge, intentions, beliefs) with a particular part of the brain.[83] Then, second, this is followed by the claim that because that particular brain part is shown (through neuroimaging) to be activated, the activated brain structures are causing the behavior or are instantiations of the faculty or attribute (often without the subject's awareness). The conceptual presuppositions underlying this picture are problematic—once exposed, they undermine many of the claims made on behalf of neuroscience in law.

A. Lie Detection

Neuroscience-based lie detection provides one of the more significant contributions neuroscience may make to law.[84] Currently, two companies are marketing fMRI-based lie detection services for litigation purposes.[85] The possibilities have garnered scholarly attention,[86] and they have caught the attention of the culture more generally.[87] The issue of neuroscience-based lie detection also stands out as a prime example in which scientists and legal scholars conflate the important distinction between criterial and inductive evidence and incoherently ascribe properties of a whole (a person) to its parts (areas of the brain).

Although important differences underlie the various neuroscience-based lie-detection studies and techniques, they share the common as-

83. One popular expression of this conception is the so-called "modular theory of mind." *See* MICHAEL S. GAZZANIGA, NATURE'S MIND: THE BIOLOGICAL ROOTS OF THINKING, EMOTIONS, SEXUALITY, LANGUAGE, AND INTELLIGENCE 124 (1992).

84. *See* Ganis et al., *supra* note 18; Kozel et al., *supra* note 7; Kozel et al., *supra* note 18; Langleben et al., *supra* note 18; Lee et al., *supra* note 18; Spence et al., *supra* note 18.

85. These companies are Cephos, http://www.cephoscorp.com (last visited May 14, 2010), and No Lie MRI, http://www.noliemri.com/ (last visited May 14, 2010). A third brain-based lie detection technique, known as "brain fingerprinting," relies on an EEG test. *See* Brain Fingerprinting Laboratories, http://www.brainwavescience.com/ (last visited May 14, 2010).

86. *See* Greely & Illes, *supra* note 2, at 385–405; Moriarty, *supra* note 14; Pardo, *supra* note 14; Sinnott-Armstrong et al., *supra* note 13.

87. *See* Scott T. Grafton et al., *Brain Scans Go Legal*, SCI. AM. MIND, Dec. 2006–Jan. 2007, at 30 (2006); Rosen, *supra* note 1.

sumption that deception involves stable and detectable neurological correlates.[88] As with traditional polygraphs, the neuroscience research is looking for a correlation between deceptive behavior and *something else*. With polygraphs it is increased heart rates, breathing, and perspiring;[89] with neuroscience it is increased blood flow to certain regions of the brain.[90]

The claims made regarding neuroscience-based lie detection may be conceptual or empirical. The empirical claims concern the results of various experiments designed to detect brain activity correlated with lying. These experiments typically involve people told to perform simple tasks like lying about biographical facts,[91] memorizing and occasionally lying about playing cards,[92] or "stealing" one of two objects in a room,[93] with researchers looking for similarities between people when lying or trying to detect when they are lying. The empirical claims following the experiments concern the similarities among subjects or the success in detecting when individual subjects are lying.[94] Although these claims currently face significant empirical limitations[95]—limitations that ought to prevent current use in real-life litigation settings[96]—that is not our primary concern in this Article.

Our concern is with the problematic character of the conceptual claims. While it is possible that some specific brain activity may be necessary to engage in deception (no less than to engage in sincere discourse)—and neuroscience may possibly provide good inductive evidence of that activity—it is a conceptual mistake to identify deception with that neural activity. Yet, neural activity is repeatedly identified with

88. *See* sources cited *supra* note 84. If not, there would be no point of the experiments.

89. *See* NAT'L RESEARCH COUNCIL, THE POLYGRAPH AND LIE DETECTION 12–17 (2003).

90. *See* sources cited *supra* note 84.

91. *See* Lee et al., *supra* note 18, at 159.

92. *See* Langleben et al., *supra* note 18, at 729.

93. *See* Kozel, *supra* note 7, at 606.

94. *See* Ganis et al., *supra* note 18, at 830 ("fMRI revealed that well-rehearsed lies that fit into a coherent story elicit more activation in right anterior frontal cortices than spontaneous lies that do not fit into a story, whereas the opposite pattern occurs in the anterior cingulate and in posterior visual cortex."); Kozel et al., *supra* note 7, at 611 ("We have shown that fMRI can be used to detect deception within a cooperative individual."); Langleben et al., *supra* note 18, at 727 ("Increased activity in the anterior cingulate cortex (ACC), the superior frontal gyrus (SFG), and the left premotor, motor, and anterior parietal cortex was specifically associated with deceptive responses."); Lee et al., *supra* note 18, at 161 ("Our imaging data revealed four principle regions of brain activation: prefrontal and frontal, parietal, temporal, and sub-cortical."); Spence et al., *supra* note 18, at 169 ("Attempted deception is associated with activation of executive brain regions (particularly prefrontal and anterior cingulate cortices) . . . .").

95. These limitations include (1) the narrow and highly artificial nature of the experimental situations, making extensions to real-life litigation situations unwarranted; (2) the individual differences in brain areas in the experiments themselves; and (3) that many of the areas identified in the studies have been associated with cognitive activities other than lying. For further discussion of these limitations see Greely & Illes, *supra* note 2, at 402–05; Monteleone et al., *supra* note 13, at 536–37.

96. To be admissible, the evidence would have to be shown to be applied reliably to the particular facts of each case. *See* FED. R. EVID. 702. Given the differences between the artificial nature of the low-stakes experimental settings and the more complicated, high-stakes settings of criminal and civil litigation, it is not clear at this time how reliability in the particular case could be shown.

deception in the legal literature.[97] These identifications occur primarily in one of two ways. Sometimes areas of the brain are said to be deciding when and whether to lie and then engaging in the processes to carry out this decision. For example, in summarizing many of the neuroscience studies on deception, Charles Keckler asserts that

> there was some agreement among all investigations to date that: (1) some form of "executive" function that deals with conflicting pressures, generally the anterior cingulate gyrus, was used to handle the "choice" of whether and when to lie, and (2) this often acted with some form of inhibitory mechanism to suppress the truthful response.[98]

Second, in addition to deciding when and whether to lie, neural activity is commonly identified with deception more indirectly when neuroimaging is claimed to reveal what knowledge, information, or intentions are "housed" or "stored" or "encoded" in the person's brain.[99] For example, in discussing EEG-based lie detection, Farwell and Smith assert that "the brain of the criminal is always there, recording the events, in some ways like a video camera."[100]

The criterial evidence for deception, however, is behavioral and not neurological (whether in the form of the brain "choosing" to lie or "housing" knowledge and intentions). To lie typically requires an intent to deceive. Deception involves knowing (or at least believing something to be) the truth and saying or implying the opposite, and it involves judgments about the beliefs and knowledge of the audience.[101] At most, neuroscientific evidence might be able to provide well-grounded empirical correlations between this behavior and brain states.[102] This will be *in-*

97. *See infra* notes 98–100 and accompanying text.

98. Keckler, *supra* note 7, at 535; *see also* Ganis et al., *supra* note 18, at 830 (claiming that neuroscience can "examine directly the organ that produces lies, the brain"); Wolpe, Foster & Langleben, *supra* note 15, at 39–40 (noting the possibility that neuroimaging will allow scientists to "peer into an individual's thought processes").

99. *See* Denno, *supra* note 6, at 333 ("[EEG lie detection] is based upon the principle that the human brain houses information . . . ."); Keckler, *supra* note 7, at 510 (claiming that neuroscience may be used to "accurately distinguish between the presence and absence of" knowledge in a person's brain); *see also* John-Dylan Haynes et al., *Reading Hidden Intentions in the Human Brain*, 17 CURRENT BIOLOGY 323 (2007).

100. Farwell & Smith, *supra* note 18, at 135; *see also* Andre A. Moenssens, *Brain Fingerprinting—Can It Be Used to Detect the Innocence of Persons Charged with a Crime?*, 70 UMKC L. REV. 891, 903 (2002) ("[EEG-lie detection] at its best, can only detect whether certain knowledge exists in the subject's brain.").

101. The relationship between deception and lying is itself complicated: one can lie without deceiving (or even intending to deceive) and one can deceive (or intend to deceive) without lying. For a discussion, see Don Fallis, *What Is Lying?*, 56 J. PHIL. 29 (2009). The conceptual points regarding deception described in the text above are acknowledged in many of the same studies that then go on to identify lies with brain states. *See, e.g.*, Kozel et al., *supra* note 7, at 605 ("Lying is a complex process requiring suppression of the truth, communication of a coherent falsehood, . . . and modifications of behaviors to convince the receiver of ones actions."); Lee et al., *supra* note 18, at 163 ("[A]n essence of lying is the recognition of, and attempt to manipulate, the mental states of others."); Spence et al., *supra* note 18, at 172 ("Deceiving another human subject is likely to involve multiple cognitive processes, including theory of mind concerning the victim's thoughts (their ongoing beliefs) . . . .").

102. *But see* Monteleone et al., *supra* note 13, at 536–37.

ductive evidence of such behavior. The neuroscience evidence, in other words, may be able to provide a *measurement* of deception, but not the *measure* of it.[103] It is a conceptual mistake, therefore, to conclude that lies take place in the brain (when a particular area of the brain chooses to lie); that neuroscience can reveal lies being "produced" in the brain; or that it can reveal whether knowledge, information, or intentions are located in someone's brain. If there is a discrepancy between the behavior associated with deception and previously identified brain states, the behavioral evidence will override the inductive (neuroscientific) evidence.

To illustrate this, suppose the reverse were true. If particular brain states *did* provide criterial and not just inductive evidence of deception, then by hypothesis having certain neural activity would be a sufficient condition for engaging in an act of deception—even if one did not intend to deceive and one uttered what she thought to be true. Would we really say this person was lying? Of course not, and the example makes plain, first, that what constitutes "deception" or a "lie" is a conceptual not an empirical question,[104] and, second, that the criteria are behavioral not neurological. Certain brain states may be causally *necessary* for deception, but they are not a sufficient condition for deception.

The confusion of criterial with inductive evidence in the lie-detection context thus leads directly to confusing properties of wholes (persons) with properties of their parts (brains and brain regions). This occurs whenever properties that may intelligently be ascribed only to the person as a whole (e.g., engaging in deception) are instead ascribed to one of its parts (the brain). Moreover, the fallacy has important practical, not only theoretical, implications in this context because the evidential value of neuroscientific evidence is affected by how it is conceived. If lies, knowledge, and intent are mistakenly thought to be identical with particular brain states, then proof of the brain states will (again, mistakenly) appear to provide conclusive proof of the corresponding mental states.[105] Indeed, those in the grip of this fallacy might even suggest that we attach legal consequences directly to people who have brains in particular neurological states (eliminating the "mental middle-man," as it were). Once it is recognized, however, that such evidence may have some inductive correlation with (and may even be causally necessary for) certain behavior—but does not provide a sufficient condition to establish

103. The behavior may provide both.

104. To further illustrate this, imagine trying to construct an experiment to prove whether lies are really brain states *and not* situations in which someone says something false to another.

105. A recent conviction in India based on use of brain-scan lie-detection evidence illustrates how mistaken conceptual assumptions have significant practical consequences. State v. Sharma, No. 508/07 (Session Ct. of Pune June 12, 2008) (India), http://court.mah.nic.in/courtweb/orders/pune/pundcis/orders/201501005082007_1.pdf; *see* Anand Giridharadas, *India's Novel Use of Brain Scans in Courts is Debated*, N.Y. TIMES, Sept. 15, 2008, at A10 ("But it was only in June, in a murder case in Pune, in Maharashtra State, that a judge explicitly cited a scan as proof that the suspect's brain held 'experiential knowledge' about the crime that only the killer could possess, sentencing her to life in prison.").

the categories at issue—then the value of the evidence diminishes accordingly.[106]

B. Criminal Law

As with lie detection, aspects of criminal law doctrine have attracted a great deal of attention from the neurolaw community. At stake are issues involving voluntary conduct (*actus reus*), mental states (*mens rea*), diminished capacity, insanity, theories of punishment, and the death penalty.[107] Our discussion focuses primarily on the first two issues, voluntary action and mental states.[108] Scholars have made strong claims about the ways neuroscience can and ought to, or potentially will be able to, aid our understanding of these fundamental doctrinal issues.[109] These claims typically confuse properties of people with properties of brains.

Criminal liability typically requires conduct that includes a voluntary act by the defendant. Modern criminal statutes, including the Model Penal Code, exclude from the category of voluntary acts bodily movements such as reflexes, convulsions, sleepwalking, acts performed while under hypnosis, and other movements that are not "a product of the effort or determination of the actor."[110] Professor Deborah Denno challenges this dichotomy between voluntary and involuntary acts, arguing that "the new neuroscientific research on consciousness"[111] has "confirm[ed] that there appears to be no sound scientific basis" for the distinction.[112]

Professor Denno characterizes voluntary acts as involving three components: "(1) an internal event, or volition; (2) an external, physical demonstration of that volition; and (3) a causal connection between the

106. This is not to suggest that it might not be probative in certain cases; it just will not be conclusive. *See* Pardo, *supra* note 14, at 315–18; *see also* Moriarty, *supra* note 14, at 47–49; Sinnott-Armstrong et al., *supra* note 13, at 362–67. Deception may be similar to pain in the sense that whereas certain behavior provides criterial evidence for pain, strong inductive correlations between cortical activity and being in pain suggests that the cortical activity may provide probative evidence of pain. On the neuroscience of pain, see Kolber, *supra* note 7.

107. *See* Buckholtz et al., *supra* note 6, at 934–36 (discussing third-party punishment); Denno, *supra* note 6 (discussing involuntary conduct); Greene & Cohen, *supra* note 8, at 209–10 (discussing theories of punishment); Stephen J. Morse, *Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience*, 9 MINN. J. L. SCI. & TECH. 1, 13–19 (2008) [hereinafter Morse, *Determinism*] (discussing free will and criminal responsibility); O'Hara, *supra* note 6, at 27–30 (discussing mental states); Snead, *supra* note 11 (discussing capital punishment). Professor Morse has been a trenchant, and in our view largely successful, critic of many of the claims made on behalf of neuroscience regarding criminal responsibility. *See* Stephen J. Morse, *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*, 3 OHIO ST. J. CRIM. L. 397 (2006) [hereinafter Morse, *Brain Overclaim Syndrome*]; Stephen J. Morse, *New Neuroscience, Old Problems*, *in* NEUROSCIENCE AND THE LAW: BRAIN, MIND, AND THE SCALES OF JUSTICE 157 (Brent Garland ed., 2004) [hereinafter Morse, *New Neuroscience*].

108. We return to the issue of voluntary actions in Part IV *infra*.

109. *See, e.g.*, O'Hara, *supra* note 6, at 27–31.

110. MODEL PENAL CODE § 2.01 (1962).

111. Denno, *supra* note 6, at 320.

112. *Id.* at 328.

internal and external elements."[113] She states that willed movements "lie at the heart" of voluntary acts—"[i]n other words, when do people consciously feel they have engaged in a voluntary act?"[114] Although willed movements are at the heart of voluntary acts, she contends that "[s]ome of the most powerful research in neuroscience suggests that the unconscious may be in charge of how human beings make decisions about willed movements."[115] The "powerful research" is the work of Benjamin Libet,[116] which she describes as follows:

> A typical Libet experiment—much simplified for this discussion—goes as follows: Libet would ask human subjects to make hand movements whenever they felt like it while he measured the electrical activity in their brains. With EEG recordings, this could be done with almost millisecond precision. Libet found that the subjects' brain impulses associated with their movements began about 300 to 350 milliseconds—or about a third of a second—before the subjects reported any conscious awareness of their intention to make the movement. In essence, the motor-planning areas in their brains began to stir a third of a second prior to when the subjects became aware of the desire to act. According to Libet and others, a subject's decision to move a finger or a wrist must have originated unconsciously and only appeared to that person as a conscious wish about a third of a second later.[117]

According to Denno, the "now accepted" view of consciousness based on this research is that "consciousness evolves gradually, starting from the unconscious and moving toward to pre-conscious states on the way to becoming a settled state of consciousness. What seems like two modes of processing [conscious and unconscious] is really a whole brain reaction."[118] Libet's research purports to "confirm[] that there appears to be no sound scientific basis" for the criminal law's distinction between voluntary and involuntary conduct.[119]

113. *Id.* at 275–76.

114. *Id.* at 326.

115. *Id.*

116. *See, e.g.*, Benjamin Libet, *Are the Mental Experiences of Will and Self-Control Significant for the Performance of a Voluntary Act?*, 10 BEHAV. & BRAIN SCI. 783 (1987); Benjamin Libet, *Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action*, 8 BEHAV. & BRAIN SCI. 529 (1985). For a critique of Libet, see Morse, *Determinism*, *supra* note 107, at 29–31.

117. Denno, *supra* note 6, at 326–27 (footnotes omitted). Denno also suggests that the movements in the experiments may be similar to firing a gun. *Id.* at 326. *But see* Morse, *Determinism*, *supra* note 107, at 30–31 ("Libet's task involved 'random' finger movements that involved no deliberation whatsoever and no rational motivation for the specific movements involved. This is a far cry from the behavioral concerns of the criminal law or morality, which address intentional conduct in contexts when there is always good reason to refrain from harming another or to act beneficently. In fact, it is at present an open question whether Libet's paradigm is representative of intentional actions in general because Libet used such trivial behavior." (footnotes omitted)).

118. Denno, *supra* note 6, at 328. Denno notes that the conscious mind may still have a "veto" option in the "150 to 200 milliseconds" between when the actor becomes aware of the intention and the act. *Id.* at 327. This would reduce voluntary action to something akin to deciding whether to inhibit a sneeze or a cough that you feel is about to occur.

119. *Id.* at 328.

It is incoherent to ascribe voluntary conduct to brain processes. The initial conceptual mistake is to associate voluntary action with an "internal process" or "feeling" of consciously engaging in voluntary conduct.[120] Having a prior, internal "feeling" is neither a necessary nor a sufficient condition for voluntary acts.[121] One may move voluntarily—for example, in answering the telephone when it rings, or typing this sentence—without first feeling an urge, desire, or intention to act. Moreover, feeling an urge to sneeze or cough just prior to doing so is not sufficient to make the subsequent sneezing or coughing voluntary. And it is a short—but mistaken—inferential leap from the association of doing something voluntarily with an internal feeling or process to associating it with the non-conscious brain activity that precedes the feeling or process, thereby equating willed movement with a "choice" made by non-conscious brain activity. Whether someone acts voluntarily does not turn on whether or not brain activity preceded a feeling or internal process; acts are voluntary when they are under the agent's control.[122]

Neurolaw claims regarding mental states provide a second example of the fallacy. As with voluntary acts, acting with a particular mental state is typically a necessary element of criminal liability. The Model Penal Code, for example, determines criminal liability based on one of four mental states: acting (1) with purpose or intent, (2) with knowledge, (3) recklessly, or (4) negligently.[123] Erin O'Hara has suggested that neuroscientific evidence may allow us to determine the first two (intent and knowledge), and to a lesser extent the third category (recklessness), because they each require the defendant to be "consciously aware of his actions and/or their harmfulness at the time of acting."[124] Professor Denno concurs, arguing that neuroscientific evidence may challenge our understandings of the purpose/intent and knowledge requirements.[125]

It is another instance of the mereological fallacy to identify either intent or knowledge with neural activity.[126] First, it is important to note that the particular doctrine largely tracks our more general concepts of

120. If there were, it would lead to an infinite regress: one would to have willed that feeling with a prior act, which would have its own feeling or inner process, and so on. This sort of mistake is also what allows determinism to get a grip. *See* Greene & Cohen, *supra* note 8, at 210–12; Morse, *Determinism*, *supra* note 107.

121. This is a conceptual point. For further elaboration, see BENNETT & HACKER, *supra* note 27, at 228–31.

122. For further elaboration of this conceptual point, see *id.* at 225–28. Moreover, as Professor Morse notes, we should not find it at all surprising that voluntary actions would be preceded by activity in the brain: "Prior electrical activity does not mean that intentionality played no causal role. Electrical activity in the brain is precisely that: electrical activity in the brain and not a mental state such as a decision or an intention. A readiness potential is not a decision." Morse, *Determinism*, *supra* note 107, at 30.

123. MODEL PENAL CODE § 2.02 (1962).

124. O'Hara, *supra* note 6, at 29.

125. Denno, *supra* note 6, at 386–87.

126. Even if it were, a further limitation would be that we typically would not get to test the defendants' brains while they are committing the crimes. But that is neither here nor there for our purposes.

intent and knowledge. Under the Model Penal Code, for example, one acts "purposely" (that is, with intent) when it is his "conscious object" to engage in conduct or to cause a particular result, and one acts "knowingly" when she is aware that her conduct is of a particular nature or that certain circumstances exist (or she is "practically certain" that her conduct will cause a result).[127] Exactly how neuroscience is supposed to inform these issues is not entirely clear. One possibility is to ask the defendant and see (via neuroscience-based lie detection) whether he is engaged in deception.[128] Regardless of the details of how neuroscience is meant to inform our understanding of intent or knowledge, in principle, brain states do not provide the appropriate criteria for either.[129]

Consider first, intention (or purpose). A person engages in conduct intentionally, or purposefully, when he knows he is engaged in that conduct and is acting for some reason (even if the reason is he just wanted to).[130] A voluntarily action need not be intentional, and intentional action need not be voluntary (e.g., when under duress).[131] From the first-person perspective, intentions may or may not be formed by previous, explicit deliberation or reasoning (it may simply be manifest in conduct—as when I respond to the telephone ring). From the third-person perspective, the person's linguistic and other behavior provide (defeasible) criterial evidence of their intentions. The upshot here is that intentions are not inner brain processes or feelings, nor are they the neural activity that preceded internal processes or feelings.[132]

Now, consider knowledge. As discussed above and, as with intention,[133] the criteria for the ascription of knowledge are various kinds of

127. MODEL PENAL CODE § 2.02(2)(a)–(b).

128. O'Hara provides the example of using neuroscience to separate deception (and thus fraud) from self-deception (and thus no fraud). O'Hara, *supra* note 6, at 28–29; *see also* Aharoni et al., *supra* note 6, at 151–53 (discussing possibilities for measuring *mens rea* indirectly with neuroscience).

129. *See* Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV. 463, 473–75 (1992).

130. This ability is conceptually distinct from one's brain states, although having a properly working brain is necessary to engage in this conduct. One's intentions may be located by finding the appropriate descriptions that explain why one engaged in the conduct. *See* G.E.M. ANSCOMBE, INTENTION 11–12 (2d ed. 1963). To illustrate this point, consider Anscombe's famous example: the same act may be described as (1) contracting one's muscles, (2) moving one's arm up and down, (3) pumping water, (4) replenishing the house's water supply, and (5) poisoning its inhabitants. *Id.* at 37–45. The descriptions that supply motivational significance for the actor help to locate her intentions. *See also* MICHAEL E. BRATMAN, FACES OF INTENTION (1999); R.A. DUFF, INTENTION, AGENCY AND CRIMINAL LIABILITY 47–51 (1990); Kimberly Kessler Ferzan, *Beyond Intention*, 29 CARDOZO L. REV. 1147 (2008).

131. But one cannot at the same time act intentionally and involuntarily. *See* P.M.S. HACKER, HUMAN NATURE: THE CATEGORICAL FRAMEWORK 212 (2007).

132. To further illustrate this last point, consider someone's statement "I intend to X, but X is impossible." It is not clear what this person even means. But suppose having an intention *just was* having a particular brain state or neural activity. Now, we would understand what the same person meant when he said, "I have a brain in a certain neural state (= the intention X), but X is impossible." Because one statement makes sense and the other does not, the brain state and the intention must not be the same thing.

133. For a similar analysis with respect to "understanding," see G.P. BAKER & P.M.S. HACKER, LANGUAGE, SENSE AND NONSENSE 346–56 (1984).

linguistic and other behavior, not brain states. Knowledge is manifested, for example, by asserting true propositions and evidence, identifying and correcting errors, and acting appropriately (intentionally) based on what is known.[134] It is, therefore, incoherent to suppose that knowledge is "housed" or "located" in the brain.[135] As with lies or deception, neuroscience may provide inductive evidence of, and possibly even identify necessary conditions for, intent or knowledge, but it cannot provide sufficient conditions for either.[136]

C. Neuroeconomics

Neuroscientific approaches to economic decision making have also emerged,[137] generating both scholarly attention in law[138] and general attention in the media.[139] The field of neuroeconomics provides particularly vivid examples of misascribing psychological attributes to the brain. The field investigates the neurological activity of people while engaged in economic decision making. Neuroeconomics is akin to the field of behavioral economics, which studies the extent to which, and attempts to explain why (typically, at the psychological level), people deviate from the conduct predicted by classical economic, rational-actor models.[140] Rather than at the psychological level, however, neuroeconomists seek to ex-

134. For these reasons, "to know" is characterized as a "success" or "achievement" verb in the sense that it implies that some sort of goal or function has been accomplished. *See* ALVIN I. GOLDMAN, KNOWLEDGE IN A SOCIAL WORLD 60 (1999); Dennis Patterson, *Fashionable Nonsense*, 81 TEX. L. REV. 841, 890–92 (2003) (book review); *see also supra* note 81.

135. *See, e.g.*, Denno, *supra* note 6, at 333 ("Brain fingerprinting is based upon the principle that the human brain houses information . . . ."); Farwell & Smith, *supra* note 18, at 135 ("[R]ecent advances in neuroscience allow scientists to detect information stored in the brain . . . ."). As with intentions, although it is incoherent to say, "I know X, but X is false," it is coherent to say, "I have a brain that is in a certain neural state (= the knowledge X), but X is false." Thus, the knowledge and the brain state must not be the same. Also, the concept of knowledge exists in a web of epistemic concepts along with belief, conviction, suspicion, supposition, conjecture, doubt, certainty, evidence, truth, probability, reasons, justification, and confirmation. Elucidation of the conceptual relationships between these concepts cannot be accomplished by telling us empirical information about the brain. This is a task for conceptual epistemology.

136. To do so eliminates the social, normative aspect of both. In particular, attributing knowledge or intentions to another (or oneself) involves attributing further commitments and entitlements to them (or oneself). *See* ROBERT B. BRANDOM, MAKING IT EXPLICIT: REASONING, REPRESENTING, AND DISCURSIVE COMMITMENT 213–21, 253–62 (1994).

137. Sanfey et al., *Ultimatum*, *supra* note 20, at 1755; Sanfey et al., *Neuroeconomics*, *supra* note 20, at 108; *see also* Ariel Rubinstein, *Comments on Neuroeconomics*, 24 ECON. & PHIL. 485, 485 (2008) ("Neuroeconomics will remain a hot topic in economics during the coming decade, probably one of the hottest."); Steven G. Sapra & Paul J. Zak, Neurofinance: Bridging Psychology, Neurology, and Investor Behavior (Dec. 1, 2008) (unpublished manuscript), *available at* http://ssrn.com/abstract=1323051 (discussing empirical findings in neuroeconomics).

138. *See* Chorvat & McCabe, *supra* note 3; Chorvat & McCabe, *supra* note 8; Hoffman, *supra* note 20; Purdy, *supra* note 20; Zak, *supra* note 20.

139. *See, e.g.*, Sandra Blakeslee, *Brain Experts Now Follow the Money*, N.Y. TIMES, June 17, 2003, at F1.

140. *See generally* Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998).

plain such conduct at the level of brain activity.[141] Legal scholars have then drawn on these findings in an attempt to derive further implications for the law.[142] According to one scholar, "[t]he tantalizing promise of neuroeconomics" is to render "literally visible the activity of other minds. It will let us see reason, fear, and principle at work, let us watch utility accumulate or dissipate."[143]

The current research is more modest in aim. Through fMRI, researchers monitor the brains of people while they make simple economic decisions. One example involves a neuroscience study of the "ultimatum game."[144] In this game, two participants are told that a particular sum of money is to be divided between them; player 1 will propose a division and then player 2 will choose to accept or reject it.[145] It is a one-shot game; if player 2 rejects the offer, they both walk away with nothing.[146] According to the rational-actor model, player 1 ought to propose the smallest unit above zero for player 2 and then keep the rest for himself (because this maximizes player 1's share), and player 2 ought to accept any amount offered greater than zero (because anything is greater than zero, 2's only other option).[147] Not surprisingly, people deviate routinely from this outcome. For example, in most studies, about half of the player 2s who perceived an offer as unfair rejected it.[148]

Neuroscience purports to tell us why. Researchers used fMRI to examine the brains of players presented with unfair offers; they noticed increased activity in several brain regions.[149] Three areas in particular were the "bilateral anterior insula" (which has been associated with "negative emotional states"), the "dorsolateral prefrontal cortex" (which has been "linked to cognitive processes such as goal maintenance and executive control"), and the "anterior cingulate cortex" (which has been "implicated in detection of cognitive conflict" such as between "cognitive and emotional motivations").[150] Those with greater increases in the emotional part tended to reject the unfair offers, and those with greater increases in the cognitive ("rational") part tended to accept them.[151] What exactly does this illuminate? And what implications does it have for law?

According to Terrence Chorvat and Kevin McCabe, this evidence supports a particular conception of economic (and, more generally, hu-

141. Sanfey et al., *Neuroeconomics*, *supra* note 20, at 108–09; *see also* Chorvat & McCabe, *supra* note 8, at 1242 ("Carried to their logical extreme, for example, these models might reveal that the reason a particular decision was made was a change in the membrane permeability in certain neuronal and glial cells.").

142. Chorvat & McCabe, *supra* note 3, at 124–28; Zak, *supra* note 20, at 149–50.

143. Purdy, *supra* note 20, at 39–40.

144. Sanfey et al., *Ultimatum*, *supra* note 20, at 1755.

145. *Id.*

146. *Id.*

147. *Id.*

148. *Id.*

149. *Id.* at 1756.

150. *Id.* at 1756–57.

151. *Id.* at 1757–58.

man) decision making as the product of different brain processes in competition with one another.[152] The "key questions" to be asked are "how does the brain decide which problems it will address"? and "what neural mechanisms are used to solve the problem?"[153] With regard to the ultimatum game, different brain regions (the emotional and cognitive) "seem to embody different thought processes."[154] In addition, in both those who accepted and those who rejected unfair offers, the "anterior cingulate cortex," which is "clearly involved in cognitive conflict resolution,"[155] was also "significantly active."[156] In the ultimatum game, the anterior cingulate cortex therefore "seems to moderate between these different regions."[157]

To summarize, the neuroeconomics interpretation of what occurred for each subject goes something like the following: Subject is faced with an unfair offer. Subject's brain must decide what to do. Subject's brain must decide which process to use to decide this question. Two brain processes begin to analyze the problem—the emotional[158] and the cognitive. They may reach contradictory conclusions. If so, a third part of the brain then adjudicates between them, deciding to go with one or the other.

It is not clear how seriously we are supposed to take these characterizations, but this summary is not an exaggeration. For instance, the researchers themselves define the issues for future research as "[u]nder what circumstances do these various systems [in the brain] cooperate or compete? When there is competition, how and where is it adjudicated?"[159] The ultimatum game is just one of several examples in the neuroeconomics decision-making literature.[160] In each, different patterns of brain activation during a decision-making task is interpreted to mean that different brain processes are working, sometimes together and sometimes in competition, with perhaps some other part of the brain adjudicating between them, deciding what the brain (that is, the person)

152. Chorvat & McCabe, *supra* note 8, at 1242. The neuroeconomics literature provides an interesting case study in how the vocabulary of one discipline gets transposed into another (e.g., talk of "competition" among brain locations). *See also infra* note 169 and accompanying text (analogizing the brain to the corporation).

153. Chorvat & McCabe, *supra* note 8, at 1248.

154. *Id.* at 1253.

155. *Id.* at 1249.

156. *Id.* at 1253.

157. *Id.*

158. Emotion is also explained in highly reductionist terms (i.e., as designed to punish or reward behavior). *See id.* at 1249.

159. Sanfey et al., *Neuroeconomics*, *supra* note 20, at 114.

160. For a discussion of others and their possible implications for law, see Purdy, *supra* note 20, at 5–15, 25–40. Professor Purdy, it should be noted, appears to be one of the few writing in the neurolaw area who rejects the hard materialist/physical conception of mind, but exactly what his conception is, is not entirely clear (at least to us). For example, he writes that "seeing a brain is not the same as seeing a mind," but later he concludes that observing patterns of correlations among neural activities may "bridge some of the distance between observer and observed, diminishing the importance of the problem of other minds." *Id.* at 24, 39.

will do.[161] Given the nascent stage of the research, the legal implications that then fall out of these interpretations are vague at best.[162] Chorvat and McCabe suggest the findings may help to construct legal regulations that assure greater compliance by citizens (because less likely to trigger their emotional, "reject" response, we suppose) and greater social pressures to conform to legal norms (because of emotional responses).[163]

The evidence does not support their interpretations. First, it makes no sense to say that the brain "decides," "reasons," or "adjudicates" anything.[164] Second, all that the neuroscientific evidence shows with regard to the ultimatum game is what subjects' brains were doing while they (the subjects) were deciding whether to accept or reject the offer. Consider the following analogy. Suppose one's face turned red whenever he was angry. Now, suppose when faced with an unfair offer in the ultimatum game, his face turned red right before he rejected the offer. Surely we would not say that the person's face rejected the offer—why, then, conclude that his insula cortex did so because it too turned colors on an fMRI machine?[165]

Emotional reactions have objects and causes.[166] Moreover, people react emotionally to what they perceive as just and unjust, fair and unfair. To be sure, these reactions are affected by one's background beliefs and knowledge about what constitutes fair and unfair conduct, of how people ought to treat and be treated by one another. If so, then people may reject unfair offers because of their emotional reactions. And these reactions may have been caused by the unfair offers themselves.[167] Having a properly working brain (including a properly working insula cortex) may make it *possible* for one to have this emotional reaction, but it is the

161. *See id.* at 5–15.

162. *See* Rubinstein, *supra* note 137, at 493 (questioning the relevance of these studies for economics).

163. Chorvat & McCabe, *supra* note 3, at 127.

164. Bennett and Hacker explain,

> It is not that as a matter of fact brains do not think, hypothesize and decide, see and hear, ask and answer questions; rather, it makes no sense to ascribe such predicates *or their negations* to the brain. The brain neither sees, *nor is it blind*—just as sticks and stones are not awake, *but they are not asleep either*. The brain does not hear, but it is not deaf, any more than trees are deaf. The brain makes no decisions, but neither is it indecisive. Only what *can* decide can be indecisive. . . . *The brain is not a logically appropriate subject for psychological predicates*.

BENNETT & HACKER, *supra* note 27, at 72.

165. The neuroeconomic interpretations are similar to lie detection in this respect. Just because a part of the brain is illuminated in an fMRI test during behavior does not necessarily mean that this part is causing the behavior. Similarly, if an increased heart rate occurs when someone is lying, we would not (for good reason) say his heart is causing the lie.

166. These may, but need not, be the same thing. For example, a loud noise may be the cause of my fear, but the object may be that there is a burglar in the house. *See* BENNETT & HACKER, *supra* note 27, at 206 ("[W]hat *makes* one jealous is not the same as what one is jealous of; your indignant tirade may make me feel ashamed, but what I am ashamed of is my own misbehaviour; a change in the fortunes of war may make one feel hopeful, but what one hopes for is final victory.").

167. *Cf. id.* at 216 ("If one is indignant at a perceived injustice, what tells one that the object of one's indignation is an evil is not that one feels flushed . . . . On the contrary, one is indignant at A's action because it is unjust, not because one flushes in anger when one hears of it. And one knows it to be unjust because it rides roughshod over someone's rights, not because one flushes in anger.").

person who feels the emotion and the person (not his brain) who decides to reject the offer.

The neuroeconomics approach to explaining decision making misguidedly ascribes psychological attributes to the brain (e.g., deciding, reasoning, adjudicating) which only make sense when attributed to the whole person. But things get worse. Professor Sanfey and colleagues take things one step further by comparing the brain with organizations of people.[168] Forgive the lengthy, but illustrative, quote:

> There are striking parallels between the brain and a modern corporation. Both can be viewed as complex systems transforming inputs into outputs. Both involve the interaction of multiple, highly similar, agents (neurons are similar to one another, just as are people), which, however, are specialized to perform particular functions. Thus, in corporations, units often take the form of departments that perform functions such as research, marketing, and so on. Similarly, the brain has systems specialized for different functions. As in a corporation, these functions may be more or less spatially segregated in the brain, depending upon the processing requirements of the specific functions and their interactions.
>
> Furthermore, there is hierarchical structure in both brains and corporations. Both rely on 'executive' systems that make judgments about the relative importance of tasks and decide how to mobilize specialized capabilities to perform those tasks.[169]

For the same reasons it is a mistake to ascribe human attributes to the brain or its parts, it is a mistake to ascribe the qualities of groups of people to the brain and its parts. The analogy of the brain to the modern corporation is more confusing than illuminating.

D. Moral Decision Making

In a series of much-discussed studies, Joshua Greene and colleagues have been exploring the neurological processes underlying moral decision making.[170] The studies present people with hypothetical "moral dilemmas" and examine their brains with fMRI while selecting what they conclude to be the moral choice.[171] As with the ultimatum game, Professors Chorvat and McCabe draw implications for law.[172]

In the primary study, people were presented with two variations of the "trolley problem."[173] In the first scenario, people were told that a train traveling down a track will kill five people unless they flip a switch that diverts the train to another track, but flipping the switch will kill one

168. Sanfey et al., *Neuroeconomics*, *supra* note 20, at 109.
169. *Id.*
170. *See* sources cited *supra* note 21.
171. *See* sources cited *supra* note 21.
172. Chorvat & McCabe, *supra* note 3, at 119–20.
173. Greene et al., *An fMRI Investigation*, *supra* note 21, at 2105–06.

person on the second track.[174] Most people flipped the switch.[175] In the second scenario, people were told that they are on a footbridge above the track, with the train again heading toward five people.[176] A large person is next to them—if they push this person, the body of the pushed person will stop the train, but the pushed person will be killed.[177] Most people chose not to push.[178]

Along with the general difference in conclusions, accompanying differences occurred at the neural level of the subjects. In the footbridge scenario, areas of the brain associated with emotion were "significantly more active" than in the first scenario.[179] Decision making in the first scenario, by contrast, involved increased activation in areas of the brain associated with cognitive processing.[180] From this and subsequent studies, Greene and his colleagues posit that "personal" moral dilemmas (such as the decision to push) involve more emotional reactions than "impersonal" ones (such as the decision to flip the switch).[181] They go even further. Because the impersonal, less emotional decisions were generally consistent with utilitarian outcomes and the personal, more emotional decisions were generally consistent with deontological outcomes, they posit that different brain areas (emotional and cognitive) may control different types of moral reasoning (utilitarian and deontological).[182] Greene and his colleagues thus conclude that there is no one "moral" part of the brain; rather they conclude that "the ordinary concept of moral judgment refers to a variety of more fine-grained and disparate processes."[183]

Although Greene is cautious about the normative implications of these findings,[184] Chorvat and McCabe push on. They argue that the findings "tend to indicate" that the more impersonal a decision is, the more "rational" the decision-making process will be, and thus better at producing "socially optimal choices."[185] As an example, they suggest that jurors will make better (i.e., "socially optimal") decisions if they can keep the "subjects of the decision" impersonal ("at a distance"), and therefore

174. *Id.* at 2105.
175. *Id.*
176. *Id.*
177. *Id.*
178. *Id.*
179. *Id.* at 2107. These areas include the medial frontal gyrus, posterior cingulate gyrus, and angular gyrus. *Id.*
180. *Id.* These areas include the middle frontal gyrus and the parietal lobe. *Id.*
181. *See* sources cited *supra* note 21.
182. *See* Greene et al., *Neural Bases*, *supra* note 21, at 398.
183. Greene & Haidt, *supra* note 21, at 523.
184. *See* Greene, *supra* note 21, at 849.
185. Chorvat & McCabe, *supra* note 3, at 119–20. Regarding the experiments, they assert: "For those subjects who did decide to push the person next to them, one might argue that 'logic' or cognitive processes prevailed over 'emotion.'" *Id.* at 119.

that the law may have an interest in helping juries to keep decisions impersonal (i.e., unemotional and "rational").[186]

As with economic decision making, the descriptions of this research run into conceptual problems. To assign moral decision making to parts of the brain is misconceived. That subjects would have emotional reactions to "personal" hypothetical situations is no surprise. Nor is it surprising that situations producing strong deontological convictions (or more categorical rules regarding right and wrong) would also be ones to which subjects reacted emotionally.[187] As with the ultimatum game, it begs the question to suppose that the brain processes activated during a decision (those associated with emotion) caused the decision, rather than both the brain processes and the decision being reactions to the situation.

As for which decision is moral, right, rational, or "socially optimal," it also begs the question to suppose, as Chorvat and McCabe do, that the question can be decided by looking at which part of the brain is "activated" during decision making.[188] It is, of course, a matter of vast philosophical controversy the extent to which utilitarian or deontological considerations provide better criteria for moral judgments.[189] Fortunately we do not have to adjudicate those disputes here because, plainly, the criterion is not which part of the decision maker's brain is more active. And even if certain brain activity is inductively well correlated with both emotions and deontological judgments, on one hand, and other brain activity is inductively well correlated with utilitarian judgments, on the other, it again begs the question whether the law ought to foster the latter over the former.[190] As with the controversial questions regarding utility versus deontology, the difficult questions regarding the law's role vis-à-vis moral decision making cannot be resolved by appealing to physical states of the brain. These are not the right criteria for resolving these issues.[191]

186. *Id.* at 119–20.

187. *See supra* note 167.

188. Chorvat & McCabe, *supra* note 3, at 119.

189. *See* Greene et al., *Neural Bases*, *supra* note 21, at 398.

190. We assume, for example, that it is not obvious that the law ought to foster the killing of one healthy, innocent person in order to harvest his organs to save six people in need of transplants. More generally, the neuroscience may not be playing any non-question-begging role in debates between utilitarian and deontological moral theories. *See* Selim Berker, *The Normative Insignificance of Neuroscience*, 37 PHIL. & PUB. AFF. 293, 294 (2009) ("[E]ither attempts to derive normative implications from these neuroscientific results rely on a shoddy inference, or they appeal to substantive normative intuitions (usually about what sorts of features are or are not morally relevant) that render the neuroscientific results irrelevant to the overall argument."). For a discussion of additional complications with attempts to derive normative conclusions from neuroscientific data, see F.M. Kamm, *Neuroscience and Moral Reasoning: A Note on Recent Research*, 37 PHIL. & PUB. AFF. 330, 341–45 (2009).

191. Although different in a number of respects from the work of Greene and his colleagues, the recent work of John Mikhail on universal moral grammar also assumes—mistakenly, in our view—that moral knowledge is contained or encoded in the brain. *See* John Mikhail, *Moral Grammar and Intuitive Jurisprudence: A Formal Model of Unconscious Moral and Legal Knowledge*, *in* THE PSYCHOLOGY OF LEARNING AND MOTIVATION: MORAL JUDGMENT AND DECISION MAKING 27, 29 (Daniel M. Bartels et al. eds., 2009); John Mikhail, *Universal Moral Grammar: Theory, Evidence and the Future*, 11 TRENDS COGNITIVE SCI. 143, 143–44 (2007). For a more detailed discussion of this as-

E. Jurisprudence

Along with specific legal issues and theoretical approaches, the neurolaw movement also includes general jurisprudential theories of law within its scope. Oliver Goodenough, in particular, is supremely confident in the power of neuroscience to revamp our view of law. In an award-winning article,[192] Goodenough makes the case that neuroscience will dispel our Cartesian presuppositions about the nature of law and turn our attention to the role of mind in legal reasoning.[193] From our point of view, Goodenough's claims—and the arguments he makes in the service of them—well illustrate the overblown and tendentious presuppositions of neurolegalists.

Goodenough's arguments develop against the background of the intellectual history of legal theory. In the nineteenth century, Langdell aspired to a "science" of law—a "top down" approach to the explication of legal doctrine that, in essence, evinced "a form of systematic textual analysis."[194] This emphasis gave way in the era of Legal Realism, when focus shifted from doctrine to sociology.[195] The Realist emphasis on the social scientific study of law is now poised to make a great leap forward with the advent of neuroscientific investigation of "how law actually works in human heads."[196]

Law, Goodenough tells us, "is a human mental activity."[197] Because it is a mental activity, we will know more about law by going "inside our heads" to see "how the human brain works."[198] It is the theory of mind (or mind as brain) that must be brought into law if only to supplant the dominant view, that being the Cartesian "dualism between the physical aspects of the brain and a nonphysical, mental world of awareness and feelings."[199]

---

pect of Mikhail's work, see Dennis Patterson, *On the Conceptual and the Empirical: A Critique of John Mikhail's Cognitivism*, 73 BROOK. L. REV. 1053 (2008).

192. Goodenough, *Mapping Cortical Areas*, *supra* note 8. The article "received the Jurimetrics Research Award for proposals for research on the scientific study of law." *Id.* at 429 n.*.

193. *Id.* at 431–33.

194. *Id.* at 430.

195. *Id.* at 431.

196. *Id.*

197. *Id.*

198. *Id.* at 431–32.

199. *Id.* at 432. Goodenough quotes Descartes as representative of the current mindset in academic law:

> I must begin by observing the great difference between mind and body. Body is of its nature always double. When I consider the mind—that is, myself, insofar as I am merely a conscious being—I can distinguish no parts within myself; I understand myself to be a single and complete thing. Although the whole mind seems to be united to the whole body, yet when a foot or an arm or any other part to the body is cut off I am not aware that any subtraction has been made from the mind. Nor can the faculties of will, feeling, understanding and so on be called its parts; for it is one and the same mind that wills, feels and understands.

This approach underlies much legal scholarship. A unified intelligence guides both day-to-day behavior and the ability to judge the behavior of others. The job of the law is to supply this intelligence with clear, word-based rules, based in sound policy.

Contemporary neuroscience, Goodenough claims, "offers better tools for understanding human thought."[200] To really make progress in our understanding of human thought, Goodenough argues, we need look no further than the modular theory of the brain,[201] which Michael Gazzaniga describes thus:

> The modular organization of the human brain is now fairly well accepted. The functioning modules do have some physical instantiation, but the brain sciences are not yet able to specify the nature of the actual neural networks involved for most of them. It is clear that they operate largely outside the realm of awareness and announce their computational products to various executive systems that produce behavior or cognitive states.[202]

As with the neuroeconomics and moral-judgment examples, Goodenough appeals to different brain locations to explain different types of decision making. The central insight of the modular theory of mind is that mental processes occur in different parts of the brain. In fact, "a separation exists in the brain"[203] such that different cortical areas of the brain perform different functions. If we embrace the thesis of the modularity of mind, what jurisprudential insights will we gain? Goodenough believes that locating the functions for law and moral reasoning will be the key to greater insight into law and our thinking in law.[204] He cites a variety of authorities for the proposition that our thinking about justice occurs in one cortical area, and our rule-based application of law occurs in another.[205] Accordingly, Goodenough concludes that "[s]cience has developed tools that can be used to test the theory that justice-based thinking occurs separately from rule-based reasoning."[206] How do they work?

Id. (quoting Rene Descartes, *quoted in* RICHARD M. RESTAK, THE MODULAR BRAIN: HOW NEW DISCOVERIES IN NEUROSCIENCE ARE ANSWERING AGE-OLD QUESTIONS ABOUT MEMORY, FREE WILL, CONSCIOUSNESS, AND PERSONAL IDENTITY 11 (1994)). For additional arguments positing Cartesianism in current legal doctrine, see SUSAN EASTON, THE CASE FOR THE RIGHT TO SILENCE 217 (2d ed. 1998) (arguing that limiting the privilege against self-incrimination to testimonial evidence, and not extending it to physical evidence, reflects a commitment to dualism); Dov Fox, *The Right to Silence as Protecting Mental Control*, 42 AKRON L. REV. 763, 793–94 (2009) (positing, likewise, that the testimonial-physical evidence distinction under the Fifth Amendment depends on mind-body Cartesian dualism). For a critique of the idea that the testimonial requirement depends on Cartesianism, see Michael S. Pardo, *Self-Incrimination and the Epistemology of Testimony*, 30 CARDOZO L. REV. 1023, 1040–45 (2009); see also Keren Shapira-Ettinger, *The Conundrum of Mental States: Substantive Rules and Evidence Combined*, 28 CARDOZO L. REV. 2577, 2580–83 (2007) (arguing that criminal-law doctrine regarding states of mind "is based on the premises of Cartesian dualism").

200. Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 433.

201. *Id.* As discussed above, the neuroeconomics advocates also make use of a modular theory of mind. *See supra* Part III.C.

202. GAZZANIGA, *supra* note 83, at 124.

203. Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 436.

204. *Id.* at 441. Similarly, the neuroeconomics advocates argue that understanding how different brain regions "make decisions" will help us understand economic reasoning and decision making. Sanfey et al., *Neuroeconomics*, *supra* note 20.

205. Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 439–41.

206. *Id.* at 439.

In thinking about justice we are aided by "a nonverbal algorithm that is programmed by some mixture of genetic blueprint, cultural heritage, and personal experience."[207] By contrast, word-based systems of thought, such as law, actuate "an interpreter module."[208] In legal activities such as the drafting of contracts, statutes, and regulations, the interpreter module serves to process legal materials through "a word-based formula[, employing] the implicit structural logic of the unarticulated system in which the [legal] norm is generated."[209]

Even if we accept Goodenough's claims regarding cortical separation between justice and rule-based decision making, what follows? Suppose we could locate the precise areas in the brain where these two functions occur: what could we infer from such a discovery? As we have argued, there is no denying that one must have a brain to think, just as one must have a brain to walk—but just as it is not the brain that walks, so too it is not the brain that thinks. The important question is whether "legal thinking" is *reducible* to brain function. To the degree he considers this question at all, Goodenough begs the question.

Recall Goodenough's contrast between Langdellian legal science and the Realist critique thereof.[210] Goodenough claims that neuroscience could tell us far more about the law than either of these theories.[211] And yet, his neurological account tells us nothing about the central element of the formalist/realist divide: the nature of law. Langdellian formalism posited a conceptual space of law that reason could grasp through reflection on the necessary conditions for a given doctrinal department of law.[212] The Realist critique denied the central formalist tenet of the logical structure of law.[213] In essence, the Realist critique was that the person making a legal decision was as important as the rule in question. Goodenough's account of law contributes nothing to this debate. Even if we assume some definition/conception of "law," simply locating where "legal thinking" occurs is not a jurisprudential contribution.[214]

In arguing for the notion that moral and legal thinking are the product of embedded algorithms, Goodenough claims that this hypothesis can be empirically tested.[215] This is impossible because the hypothesis is question-begging. First, if legal thinking is grounded in or actuated by

207. *Id.*
208. *Id.* at 436.
209. *Id.* at 436.
210. *Id.* at 430–31.
211. *Id.* at 433.
212. For a discussion of Langdell's project in the context of a "scientific" approach to law, see Dennis Patterson, *Langdell's Legacy*, 90 NW. U. L. REV. 196 (1995).
213. *See* Goodenough, *Mapping Cortical Areas*, *supra* note 8, at 431.
214. Goodenough claims to the contrary: "Advances in neurosciences and other branches of behavioral biology provide new tools and the opportunity to revisit classic questions at the foundation of legal thinking." *Id.* at 430.
215. *Id.* at 440.

a hard-wired algorithm, what explains legal disagreement?[216] Second, the "existence" of such an algorithm could never be confirmed by experiment because it has no features detectable by scientific experiment.[217] These limitations are ironic because the entire point of Goodenough's claims for neuroscience is that the science of the brain will advance our understanding of law and legal reasoning, but his proposal would neither resolve important jurisprudential questions nor provide falsifiable empirical claims. In trying to solve jurisprudential problems with science, the proposal serves neither.

IV. THE "THEORY" OF FOLK PSYCHOLOGY

Like neuroreductionists generally, neurolegalists seek to reduce the explanation of all human behavior to the causal level. Believing as they do that "the mind is the brain," neurolegalists have attempted to account for mental capacities, abilities, and processes solely at the level of cortical function.[218] As an explanatory account of the nature of the human agency, this reductionism aspires to nothing less than the replacement of our talk of belief, desires, and intentions with the language of neuroscience. Because all movement is "caused," no non-causal account of human action is permitted. In fact, the aspirations of some reductionists are greater still.

For what are called "eliminativists" or "eliminative materialists," all of our ordinary talk about mental life is just another "theory," one that competes with scientific theories in an effort to accurately capture and characterize mental life.[219] Eliminativists, for example, see our verbal descriptions of pain, anger, memory, and recognition as theories (and bad ones at that) of these aspects of human behavior.[220] True to its name, eliminative materialism seeks to eliminate our everyday language for explaining mental life—the language of "folk psychology."[221] The reason folk psychology should be eliminated, they contend, is that it is a bad explanatory theory.[222]

Before the question of elimination can arise, we must consider whether folk psychology even constitutes a "theory" (scientific or otherwise)? Why might eliminativists be tempted to think so? Theories in-

216. This dilemma plagues metaphysical realist accounts of law as well. *See* Dennis M. Patterson, *Dworkin on the Semantics of Legal and Political Concepts*, 26 O.J.L.S. 545, 553 (2006).

217. Given the manifold appeals to "science" from neurolegalists, it is ironic that many of their central claims are simply not amenable to scientific verification or falsification. The idea that we are "hard-wired" or have otherwise "internalized" a moral code is a notion familiar to the arguments of Goodenough, Mikhail, and many other neurolegalists. But, as Richard Rorty argued, this central claim is simply not provable. *See* Richard Rorty, *Born to Be Good*, N.Y. TIMES, Aug. 27, 2006, at 26 (reviewing MARC D. HAUSER, MORAL MINDS (2006)); *see also* Rorty, *supra* note 67, at 222.

218. *See supra* notes 39–40 and accompanying text.

219. *See* Churchland, *supra* note 40, at 68–72; Ramsey, *supra* note 40.

220. *See* Ramsey, *supra* note 40.

221. *See* Churchland, *supra* note 40, at 76.

222. *Id.* 73–76.

volve the application of concepts. And folk psychology contains concepts such as "intention," "belief," and "desire," which people apply to themselves and others. Therefore, so the argument might go, it qualifies as a theory.

It is difficult to avoid the conclusion that this tenet of the eliminativist critique of "folk psychology"—i.e., that it is a bad "theory"—is wildly implausible. Our workaday concepts of intention, cause, and responsibility are not theories, for a scientific theory has to be amenable to verification or falsification. A network of concepts is not necessarily "a speculative assumption or theory."[223] The vocabulary with which we make sense of our mental and social life is just that: a vocabulary.

And the use of this vocabulary is not a theoretical exercise. It is behavior. When I stub my toe and utter "Ouch!," I am not making a theoretical claim about what is going on in my brain. We see another's joy in his face, anger when his fist hits the table, and fear as he cringes in horror. The fundamental error of the eliminativist is the attempt to reduce behavior to "theory" and then to argue that the theory is defective because it cannot "explain" the inner workings of person. But our everyday vocabulary was never intended to operate as such a theory, does not function as such and, thus, cannot be tested as such.

Nevertheless, according to eliminative materialism, the belief that we control our actions through our intentions is one that will be consigned to the dustbin of history in much the same way as our belief that the earth is at the center of solar system or that the earth is flat.[224] Behavior is not driven by intentions, desire, and belief; rather, behavior is the product of causal forces over which we have no control. The promise of neuroscience is to ultimately reveal "the mechanical processes that cause behaviour."[225]

How do these theoretical concerns impact a tangible problem? We now circle back to Greene and Cohen on voluntary conduct and the problem of free will. According to Greene and Cohen, the attribution of free will to ourselves is just another instance of our ongoing self-deception about our volitional powers.[226] The promise of neuroscience is to talk us out of our illusions of self-control.[227]

223. BENNETT & HACKER, *supra* note 27, at 368.

224. *See* P.M.S. Hacker, *Eliminative Materialism*, *in* WITTGENSTEIN AND CONTEMPORARY PHILOSOPHY OF MIND 60, 83–84 (Severin Schroeder ed., 2001) ("Eliminative materialism is not a serious option, since it is not a serious possibility for the study of human nature and behaviour to jettison the concepts that define its subject matter and the use of which in discourse is partly constitutive of its subjects. Not only could students of human nature not abandon these concepts and continue to study psychology, but further . . . it would thereby be shown that that creature was not a person, nor even a human being.").

225. Greene & Cohen, *supra* note 8, at 217.

226. *Id.* at 221–22.

227. For an argument to the effect that "the problem of free will" is a non-problem for law, see Stephen J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology*, 25 BEHAV. SCI. & L. 203 (2007).

With a nod to the work of Daniel Wegner, Greene and Cohen state the problem of free will this way:

> [W]e feel as if we are uncaused causers, and therefore granted a degree of independence from the deterministic flow of the universe, because we are unaware of the deterministic processes that operate in our own heads. Our actions appear to be caused by our mental states, but not by physical states of our brains, and so we imagine that we are metaphysically special, that we are non-physical causes of physical events.[228]

If we are not metaphysically special, how and why does it matter? If we are not the uncaused causers of our behavior, and "choice" is an illusion, then there can be no such thing as "responsibility." We are no more responsible for our actions than the apple that falls from the tree. Both the apple and we are simply material objects beholden to the physical laws of the universe. We are not special. Our minds do not animate physical objects (i.e., our bodies). Like other folk psychological notions, "mind" is an illusion.

The central notion in this reductionist critique of folk psychology is "cause." In explaining human action, all reductionists (including eliminative materialists) assert that human action can be accounted for in strictly physical terms.[229] When we say that one thing caused another, we usually mean that one action or event brought about another. For example, when a bowling ball hits the pins, we say that the pins fell over because they were hit by the ball. The reason the pins fell over is because they were hit by a bowling ball. One event—the pins falling over—was caused by another: the bowling ball hitting the pins. The explanation of the action of the pins falling over is completely physical and causal. Can all events be so explained?

Consider an event as simple as a person stopping her car at a red traffic light.[230] Two events together comprise the event "stopping at a red traffic light": the illumination of the red lamp and the pressure of the driver's foot on the brake. Do the light waves from the lamp cause the pressure on the brake pedal? Surely not in the way the bowling ball causes the pins to fall over. It is true that the traffic light "causes" us to stop the car. But the "cause" for stopping the car cannot be explained solely by a physical process. By itself, the red light does not "cause" us to stop (i.e., it is not in virtue of the power of the light waves that emanate from it); rather, we stop because of the status of the light in an important social convention (i.e., the red light is a reason for stopping). The light functions as a signal because we have conferred this status upon

228. Greene & Cohen, *supra* note 8, at 218–19.

229. The technical way of putting the matter is to say that behavior (the *explanandum*) can be accounted for in strictly physicalist terms (the *explanans*).

230. The discussion that follows has greatly benefited from JEFF COULTER & WES SHARROCK, BRAIN, MIND, AND HUMAN BEHAVIOR IN CONTEMPORARY COGNITIVE SCIENCE 68–81 (2007).

it. Apart from that status, the traffic light is nothing more than an illuminated bulb inside a housing (notice that *qua* physical object, nothing establishes the lamp's status as a traffic light).

The reductionist wants to eliminate the intentional element (i.e., reason for action) in the explanation of behavior. In other words, the reductionist wants all explanation to be in terms of causes and not rules. When we ask the question "Why did you stop at the red traffic light?" the answer will appeal to a traffic rule. Although the rule accounts for your stopping (i.e., the rule is the reason for stopping), a causal account of the event cannot, without more, explain what has happened.

Reasons for action are not causes in the same sense that the bowling ball is a cause.[231] The reason we stop our car at a red light cannot be explained in the manner of a bowling ball striking a set of pins. The latter event is solely a matter of physical cause while the former is a normative explanation. Unlike the pins, we choose whether or not to stop at the red light. If we fail to stop, we run the risk of sanction. The pins in a bowling alley have no such choice: they are compelled to fall over by force of the impact of the bowling ball. We are neither bowling balls nor pins. We have a choice. It is this choice that is the ground of responsibility, which cannot be accounted for in eliminativist terms.[232]

V. CONCEPTIONS OF MIND AND THE ROLE OF NEUROSCIENCE IN LAW

What neuroscience can tell us about the mind, and about the mind's relationship to law, will depend on one's conception of mind. Put more simply, before neuroscience can tell us about something, we must have some idea about what exactly it is that neuroscience is meant to illuminate (otherwise, how else can we know what to look for?). We hope our discussions below will establish this fundamental point.

What are the options? We suggest that there are three conceptions. The first conception is classic Cartesian dualism. Under this conception, the mind is thought to be some type of non-material (i.e., non-physical) entity or thing that is a part of the person and is somehow in causal interaction with the person's body.[233] It is the source and location of the person's mental life—her thoughts, beliefs, sensations, and conscious experiences. Early neuroscientists were avowed Cartesian dualists and saw their task as figuring out how the non-physical entity known as "the mind" causally interacted with the physical brain and body of a person.[234]

231. Strictly speaking, reasons cannot be efficient (i.e., mechanical) causes.

232. Anthony Kenny has likewise argued that determinism does not undermine notions of criminal responsibility. *See* ANTHONY KENNY, FREEWILL AND RESPONSIBILITY 34 (1978).

233. *See supra* note 24 and accompanying text.

234. For a discussion of the explicit Cartesianism of early neuroscientists, see BENNETT & HACKER, *supra* note 27, at 23–67.

This first conception was later repudiated by neuroscientists and is typically the one disavowed by the neurolegalists.[235] The second conception of mind is as identical with the brain. This is the conception of Greene, Cohen, Goodenough, and the other neurolegalists. Under this conception, the mind is a material (i.e., physical) part of the human being—the brain—that is distinct from but in causal interaction with the rest of the organism. The brain is the subject of the person's mental properties (the brain thinks, feels, intends, and knows) and is the location of the person's conscious experiences. As we have argued, this conception is deeply problematic. It fails to recognize the proper criteria for the attribution of many of our psychological capacities and involves incoherently ascribing psychological attributes to the brain. This second conception, while repudiating the first, keeps intact the same logical, formal structure of the mind as a kind of entity that interacts with the body (replacing one inner agent for another: the brain for the Cartesian soul).[236] The problematic structure arises, ironically, because the mind-as-brain picture is still operating in the shadow of the Cartesian conception.

What, then, of the mind? We suggest a third conception. The mind is not an entity or substance at all (whether non-physical or physical). To have a mind is to possess a certain array of rational powers exhibited in thought, feeling, and action.[237] The roots of this conception are in Aristotle.[238] Under the Aristotelian conception, the mind is not a part of the person that causally interacts with the person's body.[239] It is just the mental powers, abilities, and capacities possessed by humans.[240] Likewise, the ability to see is not a part of the eye that interacts with other parts of the physical eye.[241] Under this conception, the question of the mind's location in the body makes no sense just as the location of eyesight within the eye makes no sense.[242]

235. *See supra* note 23.

236. For further discussion of the Cartesianism implicit in this view, see BENNETT & HACKER, *supra* note 27, at 231–35; see also HACKER, *supra* note 131, at 233–56.

237. *See supra* note 29 and accompanying text.

238. BENNETT & HACKER, *supra* note 27, at 12–23; Bennett & Hacker, *supra* note 56, at 130–33; Hacker, *supra* note 131, at 254. This third conception of mind is one shared in most crucial respects by many of the twentieth century's influential philosophers including: Wittgenstein, Sellars, Quine, Davidson, Rorty, and Brandom. A similar recent conception may also be found in VINCENT DESCOMBES, THE MIND'S PROVISIONS: A CRITIQUE OF COGNITIVISM (2001). *See also* Patterson, *supra* note 134, at 879 (reviewing DESCOMBES, *supra*). Tim Van Gelder has also suggested that "computational" theories of cognition more generally are embedded in a Cartesian philosophical framework and that rejecting the Cartesian framework may also cast doubt upon computational conceptions of cognition. Tim Van Gelder, *What Might Cognition Be, if Not Computation?*, 92 J. PHIL. 345, 379–81 (1995).

239. *See* BENNETT & HACKER, *supra* note 27, at 14–15.

240. *Id.*

241. The horsepower of a car's engine is not a part of the engine, and the ability to fly is not a part of an airplane, and so on.

242. *Cf.* BENNETT & HACKER, *supra* note 27, at 46 ("[T]he question of how the mind can interact with the body is not a question that can arise for Aristotle. Within the framework of Aristotelian thought . . . the very question is as senseless as the question 'How can the shape of the table interact with the wood of the table?'") (emphasis omitted).

As with the second conception, this Aristotelian conception is also materialist/physicalist in an important sense: if you took away or changed the physical structures in the brain, the mind would go away or change as well. Under this conception, a properly working brain is necessary for having a mind, but the mind is not identical with the brain. The criteria for the ascription of mental attributes to human beings are constituted by their manifold behaviors; it is people who think, feel, intend, and know (not parts of their brains).

What does accepting this third conception mean for law? First, it in no way implies that neuroscience cannot make valuable contributions to law. Because certain structures may be necessary to exercise various capacities or to engage in certain behavior, neuroscience may contribute greatly by identifying these necessary conditions as well as by showing that because of injury or deformity a person lacks them.[243] Neuroscience may also provide good inductive evidence of various mental capabilities. For example, if certain neurological events could be shown to be empirically well-correlated with deception in situations similar to the real-life ones that give rise to litigation, then perhaps neuroscience will be able to provide evidence that significantly raises or lowers the probability that someone is engaged in deception.

Importantly, however, this third conception implies important limitations on what neuroscience can contribute to law. We have attempted to delineate these conceptual limitations in order to keep the claims made on behalf of neuroscience within the bounds of sense. We have also provided several examples of claims that transgress this boundary. In order to illuminate our mental concepts, claims made on behalf of neuroscience for law must respect the criteria for the application of these concepts. Moreover, neuroscience cannot tell us where the brain thinks, believes, knows, intends, or makes decisions. Persons as a whole are the object for the application of these concepts, the subject of these predicates. What is more, the presence of neurological activity cannot be taken as *sufficient* for the attribution of these concepts to persons.

Understanding the conceptual issues we have raised in this Article allows for greater clarity and understanding of the tremendous empirical discoveries of contemporary neuroscience. These fundamental conceptual issues are of critical importance because they inform and underlie every important question and debate about the uses to which the law can and cannot, ought and ought not, put these discoveries.

243. *See* Tallis, *supra* note 82.