



## Alabama Law Scholarly Commons

---

Working Papers

Faculty Scholarship

---

1-7-2017

### The Promise of Neuroscience for Law: 'Overclaiming' in Jurisprudence, Morality, and Economics

Michael S. Pardo

*University of Alabama - School of Law*, [mpardo@law.ua.edu](mailto:mpardo@law.ua.edu)

Follow this and additional works at: [https://scholarship.law.ua.edu/fac\\_working\\_papers](https://scholarship.law.ua.edu/fac_working_papers)

---

#### Recommended Citation

Michael S. Pardo, *The Promise of Neuroscience for Law: 'Overclaiming' in Jurisprudence, Morality, and Economics*, (2017).

Available at: [https://scholarship.law.ua.edu/fac\\_working\\_papers/612](https://scholarship.law.ua.edu/fac_working_papers/612)

This Working Paper is brought to you for free and open access by the Faculty Scholarship at Alabama Law Scholarly Commons. It has been accepted for inclusion in Working Papers by an authorized administrator of Alabama Law Scholarly Commons.

## The Promise of Neuroscience for Law: 'Overclaiming' in Jurisprudence, Morality, and Economics

Michael S. Pardo\* and Dennis Patterson\*\*

### Introduction

Claims for the relevance and importance of neuroscience for law are stronger than ever. Notwithstanding persuasive arguments that illustrate a wide degree of 'overclaiming' in the literature, new claims alleging the importance of neuroscience for law are common.<sup>1</sup> This chapter discusses three examples of overclaiming how developments in neuroscience can contribute to issues in legal theory. The first example focuses on general jurisprudential theories about the nature of law and legal reasoning. We evaluate arguments concerning how neuroscientific evidence will contribute important insights for jurisprudential debates. The second and third examples concern moral and economic decision making, respectively. We evaluate several arguments about how neuroscientific evidence will illuminate decision making in these domains and how these insights ought to be applied to issues in law and public policy.

\* Henry Upson Sims Professor of Law, University of Alabama School of Law. This chapter draws on material from Chapter Three of MICHAEL S. PARDO & DENNIS PATTERSON, *MINDS, BRAINS, AND LAW: THE CONCEPTUAL FOUNDATIONS OF LAW AND NEUROSCIENCE* (2013).

\*\* Board of Governors Professor of Law and Philosophy, Rutgers University, New Jersey, USA; Professor of Law and Chair in International Trade and Legal Philosophy, Swansea University, Wales, UK; and Professor of Law and Chair in Legal Philosophy and Legal Theory, European University Institute, Florence, Italy.

<sup>1</sup> Stephen Morse has coined the phrase 'brain overclaim syndrome' to refer to such overclaiming. Stephen J. Morse, *Lost in Translation: An Essay on Law and Neuroscience*, in *LAW AND NEUROSCIENCE* V. 13, *CURRENT LEGAL ISSUES*, No 10, 529–62 (2010). For an example of claims about the promise of neuroscience for law, see Oliver Goodenough and Micaela Tucker, who argue that law and cognitive neuroscience are 'natural partners' and that 'advances of neuroscience are proving useful in solving some perennial challenges of legal scholarship and are leading to applications in law and policy'. Oliver R. Goodenough & Micaela Tucker, *Law and Cognitive Neuroscience*, 6 *ANN. REV. L. & SOCIAL SCI.* 61, 62 (2010). They add:

While caution is appropriate in considering neurolaw approaches, the new knowledge should—and will—be put to use. Areas of special attention in current neurolaw scholarship include (a) techniques for the objective investigation of subjective states such as pain, memory, and truth-telling; (b) evidentiary issues for admitting neuroscience facts and approaches into a court proceeding; (c) free will, responsibility, moral judgment, and punishment; (d) juvenile offenders; (e) addiction; (f) mental health; (g) bias; (h) emotion; and (i) the neuroeconomics of decision making and cooperation.

*Id.* at 61.

## I. Jurisprudence

Neuroscience advocates in law include general jurisprudential theories<sup>2</sup> among the array of subjects that can benefit from an increased role for neuroscience. Oliver Goodenough, in particular, is supremely confident in the power of neuroscience to revamp our general view of the law. In an award-winning article,<sup>3</sup> Professor Goodenough argues that neuroscience will dispel our Cartesian presuppositions about the nature of law and turn our attention to the role of the brain in legal reasoning. From our point of view, Goodenough's claims—and the arguments he makes in the service of them—well illustrate the overclaiming made by some neurolaw scholars.<sup>4</sup>

Goodenough develops his argument 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'.<sup>5</sup> This emphasis gave way in the era of Legal Realism, when focus shifted from doctrine to sociology. 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'.<sup>6</sup>

Law, Goodenough tells us, 'is a mental activity': 'We do it in our heads, with our brains.'<sup>7</sup> Therefore, we will know more about law by going 'inside our heads' to see 'how the human brain works'.<sup>8</sup> It is the theory of mind (conceptualized as the mind-as-brain) that must be brought into law to supplant the dominant view, that being the Cartesian 'dualism between the physical aspects of the brain and the nonphysical, mental world of awareness and feelings'.<sup>9</sup>

<sup>2</sup> By 'general jurisprudence' we mean claims about the nature of law. Classic approaches such as natural law (e.g. Aquinas and Finnis), legal positivism (e.g. Kelsen and Hart), and interpretivism (e.g. Dworkin) are examples. The example we discuss below purports to illuminate the divide between natural law and positivism.

<sup>3</sup> Oliver R. Goodenough, *Mapping Cortical Areas Associated with Legal Reasoning and Moral Intuition*, 41 JURIMETRICS J. 429 (2001). The article received the 'Jurimetrics Research Award for proposals for research on the scientific study of law', and was selected by 'a committee of lawyers and scientists'. See *id.* at 429 note a1.

<sup>4</sup> However, we believe there is an argument Goodenough could make that is immune from our criticisms of his general jurisprudential claims. We outline this argument below.

<sup>5</sup> Goodenough, *supra* note 3, at 430.

<sup>6</sup> *Id.* at 431.

<sup>7</sup> *Id.*

<sup>8</sup> *Id.*

<sup>9</sup> *Id.* at 432. Goodenough quotes Descartes as representative of the current mind-set 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.

*Id.* (quoting Descartes quoted in RICHARD M. RESTACK, *THE MODULAR BRAIN* 11 (1994)). Goodenough adds: 'This approach underlies much legal scholarship. A unified intelligence guides both day-to-day behavior and the ability to judge the behavior of others.' *Id.* This argument makes use of a false dichotomy between Cartesianism and a neuro-reductionist position. The persuasiveness of the argument is presented as following from a rejection of Cartesianism. However rhetorically useful this may be, nothing of significance actually follows for Goodenough's positive claims from rejecting Cartesianism. For other arguments positing Cartesianism in aspects of law and legal doctrine, see SUSAN EASTON, *THE CASE FOR THE RIGHT TO SILENCE* 217 (2d ed. 1998) (arguing that limiting the privilege against self-incrimination

Contemporary neuroscience, Goodenough claims, 'offers better tools for understanding human thought'<sup>10</sup> and 'some theory of how humans think underlies any coherent argument about the law'.<sup>11</sup> To really make progress in our understanding of human thought, Goodenough argues, we need to look no further than the modular theory of the brain, 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.<sup>12</sup>

In explaining different types of decision making, Goodenough points to different brain locations. 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'<sup>13</sup> such that different cortical areas of the brain perform different functions. If we embrace the thesis of the modularity of mind, what *jurisprudential* insights<sup>14</sup> will we gain? Locating the functions for law and moral reasoning, Goodenough believes, will be the key to greater insight into law and our thinking in law.<sup>15</sup> He cites a variety of authorities for the proposition that our thinking about justice occurs in one cortical area and rule-based application of law is located in another.<sup>16</sup> 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'.<sup>17</sup> How do they work?

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: Forensic Neuroscience and 'the Spirit and History of the Fifth Amendment'*, 42 AKRON L. REV. 763 (2009) (positing, likewise, that the testimonial-physical evidence distinction under the Fifth Amendment depends on mind-body Cartesian Dualism); Karen 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'). See also Dov Fox & Alex Stein, *Dualism and Doctrine*, THIS VOLUME.

<sup>10</sup> Goodenough, *supra* note 3, at 434.

<sup>11</sup> *Id.* at 432.

<sup>12</sup> *Id.* at 434 (quoting MICHAEL S. GAZZANIGA, *NATURE'S MIND: THE BIOLOGICAL ROOTS OF THINKING, EMOTIONS, SEXUALITY, LANGUAGE, AND INTELLIGENCE* 124 (1992)).

<sup>13</sup> Goodenough, *supra* note 3, at 435.

<sup>14</sup> We emphasize that Goodenough's central claim is that neuroscience enjoys greater purchase than other theories about the nature of law. See, e.g., *id.* at 439 ('Using the new neuroscience, we can entangle some of Kelsen's apparent contradictions') and *id.* at 429 (claiming that a 'series of brain-scanning experiments' can 'help[] us understand the neurological basis of the distinction between natural and positive law'). As we detail, we think neuroscience tells us nothing about the nature of law, but it may have something to say about the nature of persons, which will have implications for law.

<sup>15</sup> Similarly, neuroeconomics advocates argue that understanding how different brain regions 'make decisions' will help us understand economic reasoning and decision making. We discuss neuroeconomics in Section III. In the next section, we discuss a similar argument regarding 'emotional' and 'rational' areas of the brain in the context of moral decision making.

<sup>16</sup> Goodenough, *supra* note 3, at 439-41.

<sup>17</sup> *Id.* at 439. The jurisprudential payoff, Goodenough argues, is that 'justice-based thinking' will tell us about natural law, and 'rule-based reasoning' will tell us about positive law. *Id.* In fusing these two issues, Goodenough conflates theories of law and theories of adjudication. Legal positivism is consistent with judges engaging in justice-based reasoning, and natural law is consistent with rule-based reasoning.

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'.<sup>18</sup> By contrast, word-based systems of thought, such as law, actuate 'an interpreter module'.<sup>19</sup> 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'.<sup>20</sup> Goodenough proposes to test his module theory of law with a series of experiments in which lawyers, nonlawyers, and law students are scanned while answering questions about hypothetical situations, to locate the brain areas associated with what Goodenough characterizes as justice-based answers and the brain areas associated with rule-based answers.<sup>21</sup>

What follows even if we accept the claims made by Goodenough regarding cortical separation between justice-based and rule-based decision making? Suppose we could locate the precise areas in the brain where, as Goodenough would have it, these two functions occur: what could we infer from such a discovery? There is no denying that one must have a brain to think, just as one must have a brain to walk. The important question is whether 'legal thinking' is explicable solely in terms of brain function. To the extent he considers this issue, Goodenough begs the question. We shall explain.

Recall Goodenough's contrast between Langdellian legal science and the Realist critique of it. Goodenough claims that neuroscience could tell us far more about the law than either of these theories. 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.<sup>22</sup> The Realist critique denied the central formalist tenet of the logical structure of law. 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—that justice-based thinking occurs in one area of the brain and rule-based thinking occurs in another area of the brain—contributes nothing to this debate.<sup>23</sup> Under any conception of 'law', simply locating where 'in the brain' legal thinking occurs is not a jurisprudential contribution to disagreements between formalists and realists or between natural lawyers and positivists.<sup>24</sup>

<sup>18</sup> *Id.*      <sup>19</sup> *Id.* at 435.      <sup>20</sup> *Id.* at 436.

<sup>21</sup> *Id.* at 439–42. He refers to a pilot study. *Id.* at 442 note 64. It is not at all clear that this would 'test' his theory. At best it might show which parts of the brain are using more oxygen than others while the agent is engaged in some activity. It would show nothing about algorithms, genetic blueprints, or cultural heritage.

<sup>22</sup> For 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).

<sup>23</sup> See Brian Leiter, *Legal Formalism and Legal Realism: What Is the Issue?*, 16 LEGAL THEORY 111 (2010). In addition to our critique of Goodenough from the perspective of jurisprudence, a different challenge to Goodenough's proposal comes from the work of John Mikhail, who posits that the justice-based decisions that Goodenough discusses are also rule-based. See JOHN MIKHAIL, *ELEMENTS OF MORAL COGNITION: RAWLS' LINGUISTIC ANALOGY AND THE COGNITIVE SCIENCE OF MORAL AND LEGAL JUDGMENTS* (2011).

<sup>24</sup> 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

Moreover, 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. This is impossible, however, because the hypothesis is question-begging. First, if legal thinking is grounded in or actuated by a hardwired algorithm, what explains legal disagreement?<sup>25</sup> Second, the existence of such an algorithm could never be confirmed by experiment because it has no features detectable by scientific experiment.<sup>26</sup> 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.

Despite these problems, Goodenough raises an issue that does connect with legal theory. Recall the American Legal Realist claim that, at least in the realm of appellate decisions,<sup>27</sup> the law is indeterminate. By 'indeterminate', we mean that controversies at the appellate level cannot be settled by existing law, thereby forcing judges to choose between plausible alternative interpretations of the law and to be creative in fashioning legally defensible solutions to hard cases.

Now, assume that Goodenough is correct in the claims he makes about law and the brain. Suppose neuroscience can really tell us something about which areas of the brain are active when legal decisions are made.<sup>28</sup> If this is so, then the implications for legal theory are clear. If some appellate cases really are indeterminate (the Realist claim), then information about the judge or judges deciding the case is of great importance.<sup>29</sup> In other words, perhaps neuroscience really can tell us something about how judges decide hard cases, and the information it provides can be used to predict future decisions better than other variables. The information would surely be useful, and it would be far more perspicuous than the usual claims one hears about 'politics' or 'personal values' doing the real work in the appellate realm.<sup>30</sup>

legal thinking.' Goodenough, *supra* note 3, at 429. See also Jennifer A. Drobac & Oliver R. Goodenough, *Exposing the Myth of Consent*, IND. HEALTH L. REV. 2015 (discussing 'neurojuridical tools').

<sup>25</sup> This dilemma plagues metaphysical realist accounts of law as well. See Dennis Patterson, *Dworkin on the Semantics of Legal and Political Concepts*, 26 OXFORD J. LEG. STUD. 545–57 (2006).

<sup>26</sup> Given the manifold appeals to science by some neuroscience advocates in law, it is ironic that several of their central claims are not amenable to empirical verification or falsification. The idea that we are 'hard-wired' or otherwise have an 'innate' moral code in our brains is a familiar feature of such arguments. But, as Richard Rorty argued, this claim is not provable. See Richard Rorty, *Born to Be Good*, N.Y. TIMES, August 27, 2006 (reviewing MARC. D. HAUSER, *MORAL MINDS* (2006)).

<sup>27</sup> Many scholars caricature the Realists as thoroughgoing indeterminists (i.e., that they believed law was indeterminate 'all the way down'). This is false. When it came to most legal questions, the Realists provided an account of law that was consistent with the tenets of legal positivism (e.g. the Rule of Recognition, primary and secondary rules, etc.). It was only in the realm of appellate decision making that the Realists believed the law was indeterminate. For an excellent discussion, see Brian Leiter, *Legal Realism and Legal Positivism Reconsidered*, in NATURALIZING JURISPRUDENCE 59, 73–79 (2007).

<sup>28</sup> Our thanks to Kim Ferzan for suggesting this point.

<sup>29</sup> Of course, this information would have to be far more detailed than what is available now, which is the measure of oxygenated blood flow in the brain (the so-called BOLD response).

<sup>30</sup> This line of inquiry would thus be a neuroscientific approach similar to the 'attitudinal model', which seeks to predict judicial decision making based on judges' perceived political 'attitudes and values'. See JEFFREY A. SEGAL & HAROLD J. SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED* (2002). Noting the 'epistemically feeble condition' of current 'predictive-explanatory' models of legal decision making, Brian Leiter outlines a number of ways in which a more robust predictive model

Goodenough does not discuss this use of neuroscience.<sup>31</sup> But we think if (and we mean *if*) it turned out that neuroscience provided insight into the mind of a jurist in the discretion-laden context of appellate decision making, a real contribution would be made.

## II. Emotion and Moral Judgments

The relationship between law and morality is incredibly complex. The two intersect in numerous ways, and these intersections have been the subject of extensive debate in legal theory. One area of intersection concerns how and the extent to which moral decision making does and should affect legal decision making. In answering these questions, legal scholars have been turning to neuroscience for insights into moral decision making, focusing in particular on a number of studies by Joshua Greene and colleagues.<sup>32</sup> We first describe the neuroscientific studies and their claimed implications for moral decision making; we next describe some of the ways in which legal scholars have relied on these studies; finally, we explain the limitations on inferences about law that may be drawn from the studies.

The neuroscientific studies are ‘scenario studies’ in which test subjects are presented with a number of short vignettes and then asked whether particular actions are ‘appropriate’ or not.<sup>33</sup> Greene et al.’s initial study—and the one principally relied on by legal scholars making use of this work—presented subjects with over forty scenarios involving moral ‘dilemmas’ (as well as several involving ‘non-moral dilemmas’).<sup>34</sup> Two of those scenarios, however, have garnered much of the discussion because of the divergent results they produced. The scenarios involve variations on the so-called

may fit with jurisprudential issues. See Brian Leiter, *Postscript to Part II: Science and Methodology in Legal Theory*, in *NATURALIZING JURISPRUDENCE*, *supra* note 27, at 183–99 (discussing the attitudinal model and its relevance to ‘naturalized’ jurisprudence). Whether neuroscience can deliver a better predictive-explanatory model of decision making is an open empirical question, but this avenue would have greater relevance to jurisprudential issues than the avenues Goodenough discusses. Such a move would parallel developments in neuroeconomics, in which scholars are looking to improve upon the psychological explanations of behavioural economics by providing neuroscientific explanations.

<sup>31</sup> In a subsequent article, Goodenough discusses other possible legal applications of his neuroscientific approach to law; he proposes studies on emotional reactions by jurors and attitudes about intellectual property. See Oliver R. Goodenough & Kristin Prehn, *A Neuroscientific Approach to Normative Judgment in Law and Justice*, in *LAW & THE BRAIN* 77 (S. Zeki & O. Goodenough eds., 2006).

<sup>32</sup> Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 *SCI.* 2105 (2001); Joshua D. Greene et al., *The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 44 *NEURON* 389 (2004); Joshua D. Greene et al., *Pushing Moral Buttons: The Interaction between Personal Force and Intention in Moral Judgment*, 111 *COGNITION* 364 (2009).

<sup>33</sup> For a general methodological discussion of such studies, see John M. Darley, *Citizens’ Assignments of Punishments for Moral Transgressions: A Case Study in the Psychology of Punishment*, 8 *OHIO ST. J. CRIM. L.* 101 (2010). Gabriel Abend has cautioned against reliance on both the types of judgments in such studies and the ‘thin’ moral concepts at issue (e.g. ‘appropriate or not’, or ‘permissible or not’) in order to draw conclusions about morality. See Gabriel Abend, *What the Science of Morality Doesn’t Say about Morality*, 43 *PHIL. SOCIAL SCI.* 157 (2013); Gabriel Abend, *Thick Concepts and the Moral Brain*, 52 *EURO. J. SOCIOLOGY* 143 (2011).

<sup>34</sup> Greene et al., *An fMRI Investigation*, *supra* note 32.

'Trolley Problem'.<sup>35</sup> One scenario tested (called 'Standard Trolley' in the Supplemental Materials to the study) is as follows:

You are at the wheel of a runaway trolley quickly approaching a fork in the tracks. On the tracks extending to the left is a group of five railway workmen. On the tracks extending to the right is a single railway workman.

If you do nothing the trolley will proceed to the left, causing the deaths of the five workmen. The only way to avoid the deaths of these workmen is to hit a switch on your dashboard that will cause the trolley to proceed to the right, causing the death of the single workman.

Is it appropriate for you to hit the switch in order to avoid the deaths of the five workmen?<sup>36</sup>

Most test subjects said, yes, that it is appropriate.<sup>37</sup>

The second scenario ('Footbridge') is as follows:

A runaway trolley is heading down the tracks toward five workmen who will be killed if the trolley proceeds on its present course. You are on a footbridge over the tracks, in between the approaching trolley and the five workmen. Next to you on this footbridge is a stranger who happens to be very large.

The only way to save the lives of the five workmen is to push this stranger off the bridge and onto the tracks below where his large body will stop the trolley. The stranger will die if you do this, but the five workmen will be saved.

Is it appropriate for you to push the stranger on to the tracks in order to save the five workmen?<sup>38</sup>

Most test subjects said, no, that it is not appropriate.<sup>39</sup>

The divergent responses for the two scenarios are thought by some to create a puzzle because each case involves a decision of whether to kill one person in order to save five, thus suggesting that the results should be similar. In the study and in subsequent papers, Greene and colleagues seek to explain the difference by linking three separate issues: (1) whether the decision is consistent with deontological or utilitarian moral considerations; (2) whether the decision is a 'personal' or 'impersonal' one; and (3) whether the decision is correlated with areas of the brain associated with 'emotion'.

The proffered explanation begins with the role of emotion. In the Footbridge scenario, areas of the brain associated with emotion were 'significantly more active' than in the Standard Trolley scenario.<sup>40</sup> The decision making in the Standard Trolley

<sup>35</sup> See Philippa Foot, *The Problem of Abortion and the Doctrine of Double Effect*, in *VIRTUES AND VICES* (2002) (originally published in 5 *OXFORD REV.* (1967)) (introducing the trolley problem); Judith Jarvis Thomson, *The Trolley Problem*, 94 *YALE L.J.* 1395 (1985) (introducing the 'bystander' version of the problem).

<sup>36</sup> See Greene et al., *An fMRI Investigation*, *supra* note 32, Supplemental Data, available at <http://www.sciencemag.org/content/293/5537/2105.abstract>.

<sup>37</sup> *Id.* at 2105. <sup>38</sup> Supplemental Data, *supra* note 36.

<sup>39</sup> Greene et al., *An fMRI Investigation*, *supra* note 32, at 2105.

<sup>40</sup> *Id.* at 2107. The areas include the medial frontal gyrus, posterior cingulate gyrus, and angular gyrus. *Id.*



scenario, by contrast, involved increased activation in areas of the brain associated with 'cognitive' processing.<sup>41</sup> They next note that the decisions that involved more emotional responses were those they labelled as 'personal' (such as Footbridge) and that those they labelled as 'impersonal' (such as Standard Trolley) produced less emotional and more cognitive processes.<sup>42</sup> After linking the personal-impersonal distinction to the emotional-cognitive distinction, the next move is to map both distinctions onto the utilitarian-deontology distinction. 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 (deontological and utilitarian).<sup>43</sup> Subsequent studies have built on these initial results and explored a variety of related issues involving emotions and moral reasoning.<sup>44</sup>

The initial study and follow-up papers had explicitly descriptive aims and were cautious about normative conclusions.<sup>45</sup> Nevertheless, Greene has since drawn more bold and wide-ranging normative conclusions about moral judgments based on the distinction he draws between emotional and cognitive processes.<sup>46</sup> He argues that the distinction undermines deontological judgments and vindicates utilitarian judgments. Deontological judgments, he argues, are produced by the 'emotional' psychological process rather than the 'cognitive' process, and utilitarian judgments are produced by the cognitive process.<sup>47</sup> The cognitive process is more likely to involve 'genuine moral reasoning', as opposed to the 'quick', 'automatic', and 'alarm-like' deontological judgments produced by emotional responses.<sup>48</sup> This, Greene argues, undermines

<sup>41</sup> *Id.* The areas include the middle frontal gyrus and the parietal lobe. *Id.*

<sup>42</sup> The study characterizes 'personal' dilemmas as those that involve actions that '(a) could reasonably be expected to lead to serious bodily harm, (b) to a particular person or a member or members of a particular group of people, (c) where this harm is not the result of deflecting an existing threat onto a different party'. *Id.*; Joshua Greene & Jonathan Haidt, *How (and Where) Does Moral Judgment Work?* 6 *TRENDS IN COG. SCI.* 517, 519 (2002). Greene later acknowledged that this distinction does not explain some of the data; see Joshua D. Greene, *The Secret Joke of Kant's Soul*, in *MORAL PSYCHOLOGY, VOL. 3: THE NEUROSCIENCE OF MORALITY: EMOTION, DISEASE, AND DEVELOPMENT* (Walter Sinnott-Armstrong ed., 2007), but whether the judgment is 'personal' (or involves physical contact) continues to be a key variable in subsequent research; see Greene et al., *Pushing Moral Buttons*, *supra* note 32.

<sup>43</sup> See Greene et al., *Neural Bases*, *supra* note 32, at 398. See also Greene & Haidt, *supra* note 42, at 523 ('the ordinary concept of moral judgment refers to a variety of more fine grained and disparate processes').

<sup>44</sup> See Greene et al., *Neural Bases*, *supra* note 32; Greene et al., *Pushing Moral Buttons*, *supra* note 32. Selim Berker notes that, as an empirical matter, all three distinctions (personal-impersonal, emotion-cognitive, deontological-utilitarian) come apart when considering other variations on the trolley problem. Selim Berker, *The Normative Insignificance of Neuroscience*, 37 *PHIL. & PUB. AFFAIRS* 293, 312 (2009).

<sup>45</sup> See Greene et al., *An fMRI Investigation*, *supra* note 32, at 2107 ('We do not claim to have shown any actions or judgments to be morally right or wrong'); Greene & Haidt, *supra* note 42; Joshua D. Greene, *From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?*, 4 *NATURE REV. NEUROSCIENCE* 847 (2003).

<sup>46</sup> Greene, *Secret Joke*, *supra* note 42; Joshua D. Greene, *Beyond Point-and-Shoot Morality: Why Cognitive (Neuro) Science Matters for Ethics*, 124 *ETHICS* 695 (2014).

<sup>47</sup> Greene, *Secret Joke*, *supra* note 42, at 50-55. Cf. Berker, *supra* note 44, at 311 ('sorting personal from impersonal moral dilemmas [is] an inadequate way of tracking the [deontological-utilitarian] distinction. To claim that characteristically deontological judgments only concern bodily harms is nothing short of preposterous.').

<sup>48</sup> Greene, *Secret Joke*, *supra* note 42, at 65.

deontology as 'a rationally coherent moral theory'; an 'attempt to reach moral conclusions on the basis of moral reasoning'; 'a school of normative moral thought'; and as reflecting any 'deep, rationally discoverable moral truths'.<sup>49</sup> Rather, deontology is characterized as merely an attempt to rationalize our emotional responses, which are based on, and may have developed evolutionarily because of, nonmoral factors. By contrast, he contends that utilitarian principles, 'while not true, provide the best available standard for public decision making'.<sup>50</sup>

Legal scholars have followed Greene down this path, drawing normative implications for aspects of the law from Greene's studies. Many of the references to the Greene studies in the legal literature cite them for the (unobjectionable) proposition that emotions play some role in moral judgments.<sup>51</sup> Most troubling from our perspective, though, is the inference that the studies show that the 'emotional', deontological judgments are incorrect or unreliable. Consider two examples. In an article discussing international criminal law, Andrew Woods relies on the studies and contends that '[h]ow moral heuristic failure occurs has been shown using fMRI scans of the brain'.<sup>52</sup> According to Woods, when subjects 'felt an emotional surge' in the Footbridge scenario, they relied on moral heuristics (e.g. 'Do no harm'), and when they did not feel this surge they engaged in utilitarian reasoning.<sup>53</sup> Woods maintains this is relevant to international criminal law because 'strong emotional intuitions may guide decision makers to outcomes that do not maximize utility'.<sup>54</sup> Similarly, Terrence Chorvat and Kevin McCabe contend that the studies are relevant to jury decision making at trial because juries will tend to make more 'rational' decisions and 'socially optimal choices when they keep the subject of the decision at a distance'.<sup>55</sup> Therefore, the law has an interest in 'depersonalizing' jury decision making.<sup>56</sup> They suggest that evidentiary rules ought to be designed with this consideration in mind.<sup>57</sup>

We resist the normative conclusions that legal scholars have drawn from the Greene studies. Before turning to the major conceptual problems that undermine these conclusions, we shall first clarify three limitations on using the results of the studies to draw conclusions about legal issues.

First, even if subjects experience strong emotional reactions to the Footbridge scenario and less so to the Standard Trolley scenario, this does not support a two-process

<sup>49</sup> *Id.* at 70–72.

<sup>50</sup> *Id.* at 77. For a critique of Greene's normative conclusions, see Richard Dean, *Does Neuroscience Undermine Deontological Theory?*, 3 *NEUROETHICS* 43 (2010).

<sup>51</sup> See, e.g., Janice Nadler, *Blaming as a Social Process: The Influence of Character and Moral Emotion on Blame*, 75 *LAW & CONTEMP. PROBS.* 1 (2012); R. George Wright, *Electoral Lies and the Broader Problems of Strict Scrutiny*, 64 *FLA. L. REV.* 759, 783 note 155 (2012); Thomas W. Merrill & Henry E. Smith, *The Morality of Property*, 48 *WM. & MARY L. REV.* 1849 (2007).

<sup>52</sup> Andrew K. Woods, *Moral Judgments & International Crimes: The Disutility of Desert*, 52 *VA. J. INT. L.* 633, 667 (2012).

<sup>53</sup> *Id.* at 668.

<sup>54</sup> *Id.* at 669. Woods ties this point to arguments about theories of criminal punishment in the international context.

<sup>55</sup> Terrence Chorvat & Kevin McCabe, *Neuroeconomics and Rationality*, 80 *CHI-KENT L. REV.* 1235, 1252 (2005).

<sup>56</sup> *Id.*

<sup>57</sup> Terrence Chorvat, Kevin McCabe & Vernon Smith, *Law and Neuroeconomics*, 13 *SUP. CT. ECON. REV.* 35, 61 (2005).

model (emotional and cognitive) of decision making. Rather than causing the decisions, the emotional reactions may instead have simply *accompanied* decisions made for moral, deontological reasons. Indeed, as John Mikhail has pointed out, virtually all of the 'personal' scenarios presented to subjects involved violent crimes and torts.<sup>58</sup> Thus, it should not be surprising (indeed, it should be expected) that subjects had (1) emotional responses, and (2) reactions to the effect that the conduct was impermissible. This does not show that subjects were not engaged in genuine moral reasoning (and instead engaged in an emotion-driven moral failure). In short, the presence of emotion neither rules out moral reasoning, nor does it specify a particular causal role for the emotions.<sup>59</sup>

Second, the relationship between emotion and the law is incredibly complex.<sup>60</sup> Therefore, *even if* the deontological judgments were caused by emotional reactions, it begs the question to assume that the law should try to depersonalize decisions and eliminate these types of judgments. This is obvious in the context of jury decision making. The very judgments that Chorvat and McCabe contend are problematic because they may lead to suboptimal decisions are ones the US Supreme Court has found to be integral to jury decision making. In criminal cases, for example, the Supreme Court has explained that evidence may serve a legitimate role at trial by implicating 'law's moral underpinnings and a juror's obligation to sit in judgment' and that the prosecution may need such evidence to show that a conviction would be 'morally reasonable'.<sup>61</sup> In civil cases, the Court has also explained that punitive damages ought to depend, in part, on judgments of reprehensibility, blame, and moral outrage by jurors towards the actions of defendants.<sup>62</sup> This is not to suggest that emotional reactions are never a problem, both for jury decision making in particular and law in general.<sup>63</sup> Our point is merely that the Greene studies are too blunt an instrument to read off any clear policy results for the proper role of emotion in law.

<sup>58</sup> John Mikhail, *Emotion, Neuroscience, and Law: A Comment on Darwin and Greene*, 3 EMOTION REV. 293 (2011).

<sup>59</sup> See Gilbert Harman, Kelby Mason, & Walter Sinnott-Armstrong, *Moral Reasoning*, in THE MORAL PSYCHOLOGY HANDBOOK 206-42 (John M. Doris ed., 2010) (outlining several possible types of moral reasoning compatible with emotional reactions); Jesse J. Prinz & Shaun Nichols, *Moral Emotions*, in THE MORAL PSYCHOLOGY HANDBOOK 111-41 (discussing the role of emotions in moral cognition). The inference from particular brain activity to particular mental functions or processes faces a number of empirical limitations. See Russell A. Poldrack, *Can Cognitive Processes Be Inferred from Neuroimaging Data?*, 10 TRENDS IN COG. SCI. 79 (2006) (discussing limitations on drawing such 'reverse inferences'). In arguing for a two-process model of moral judgments (emotional-deontological and cognitive-utilitarian), the Greene studies point to a number of differences in the time subjects take to make judgments (e.g. subjects take longer if they must engage in more reasoning or override an initial inclination). According to a recent study, nevertheless, the timing differences depend not on the difference between deontological and utilitarian judgments, but rather whether the judgments were 'intuitive' or not. See Guy Kahane et al., *The Neural Basis of Intuitive and Counterintuitive Moral Judgment*, 10 SOC. COGNITIVE & AFFECTIVE NEUROSCIENCE (2011).

<sup>60</sup> For an illuminating discussion of this complexity, see Terry A. Maroney, *The Persistent Cultural Script of Judicial Dispassion*, 99 CAL. L. REV. 629 (2011).

<sup>61</sup> See *Old Chief v. United States*, 519 U.S. 172, 187-8 (1997).

<sup>62</sup> See *Cooper Indus., Inc. v. Leatherman Tool Group, Inc.*, 532 U.S. 424 (2001). See also Todd E. Pettys, *The Emotional Juror*, 76 FORDHAM L. REV. 1609 (2007) (discussing ways in which emotions may aid and detract from accuracy at trial).

<sup>63</sup> See FED. R. EVID. 403.

Third, we note one other mismatch between the neuroscientific studies and the uses to which they are put by some legal scholars. Greene and colleagues tested a scenario involving the driver of the trolley and a scenario involving pushing someone on the track from a footbridge. It is important to clarify that in the Standard Trolley scenario the subject is *the driver* of the trolley, not a bystander. The legal scholarship citing the studies typically does not mention this fact (describing the choice as simply whether to push a button or flip a switch, leaving it ambiguous whether it is driver or bystander).<sup>64</sup> Recall, the driver scenario was the example of an *impersonal* dilemma, and the footbridge scenario was the example of a *personal* dilemma. Exactly what it would mean for the law to further ‘depersonalize’ decisions as characterized by the studies is not entirely clear. For example, any decision by a jury is surely more ‘impersonal’ than deciding to drive a trolley into and killing one or several people. This is true even for decisions to send someone to prison or voting to impose the death penalty. We can imagine some decisions by legal actors that straddle the impersonal–personal line as it is drawn by the studies (e.g. police interactions with citizens), but that distinction is inapposite to most legal decision making.

If we turn now to the more serious conceptual problems, it begs the question to presuppose that the utilitarian judgments are correct and the deontological judgments are mistaken. This is true both generally and with the particular judgments at issue in the trolley problems. As a general matter, there are intense philosophical debates between deontologists and utilitarians; to the extent legal issues depend on a choice between conflicting moral judgments, the legal issues also depend on these philosophical debates. The Greene studies do not resolve any of the contentious normative issues; indeed, as Selim Berker argues, the empirical results are irrelevant to those debates.<sup>65</sup> Any attempt to derive a normative conclusion from the neuroscientific results depends on an ‘appeal to substantive normative intuitions (usually about what sorts of features are or are not morally relevant)’, and it is this appeal, not the neuroscience, that drives the normative inferences.<sup>66</sup> Berker illustrates this by outlining

<sup>64</sup> See, e.g., Woods, *supra* note 52, at 667 (describing Standard Trolley as ‘a train is heading down a track where five people are chatting, and the only way to save them is to switch the train’s path to another track where only one man is in the way’); Chorvat & McCabe, *supra* note 55, at 1250 (describing Standard Trolley as ‘a train is coming down a track, and, if they do nothing, the train will hit a car on the track and five people will be killed, but, alternatively, if they press a button, the train will be diverted to a side track and only one person will be killed’). Some of this ambiguity may arise from Greene’s own descriptions, which also do not mention whether the person deciding whether to flip the switch is the driver or a bystander. Although the Supplemental Data, *supra* note 36, makes clear that subjects were told they are the driver, the 2001 article, by contrast, describes the problem as ‘A runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save them is to hit a switch that will turn the trolley onto an alternate set of tracks where it will kill one person instead of five.’ Greene et al., *An fMRI Investigation*, *supra* note 32.

<sup>65</sup> See Berker, *supra* note 44. Also problematic is the assumption that utilitarian and deontological principles exhaust the basis of moral judgments.

<sup>66</sup> *Id.* at 294. Francis Kamm makes a similar point in F.M. Kamm, *Neuroscience and Moral Reasoning: A Note on Recent Research*, 37 PHIL. & PUB. AFFAIRS 331 (2009). For example, Greene and the legal scholars relying on the studies assume that the presence of emotion makes moral judgments irrational. Not only is this substantive assumption highly questionable—there are often good reasons to feel certain emotions, and to not feel emotions in some contexts is itself wrong (e.g. anger at injustice, compassion for those suffering, and joy at the good fortune of loved ones). Moreover, some moral standards involve emotional dispositions (e.g. to feel guilty when ignoring one’s duties and obligations).

what he calls 'the best-case scenario' for a normative role for neuroscience in moral judgment—suppose 'a portion of the brain which lights up whenever we make a certain sort of obvious, egregious error in mathematical or logical reasoning also lights up whenever we have a certain moral intuition'.<sup>67</sup> He asks whether we should abandon moral judgments based on these intuitions, concluding that it likely will depend on further details. If there is no connection between the two, then there does not appear to be any reason to abandon the moral intuitions. (For instance, we would not suddenly conclude that murder is not wrong because of such a link.) By contrast, if the moral judgments depend on the same sort of error or mistake present in the mathematical or logical reasoning, then we should also come to see the moral judgments as mistaken or erroneous. But, if so, then it will be the common error or mistake that undermines the judgments, not the neuroscientific results. In sum, whether the law should foster or limit particular deontological or utilitarian judgments requires a normative argument, not appeal to the areas of the brain that are active during those judgments. Areas of the brain do not provide the right criteria for resolving philosophical debates about moral judgments and the legal issues that build upon them.

With regard to the specifics of the trolley problems, it also begs the question to assume that particular judgments about the scenarios are correct or incorrect. We acknowledge the possibility that the judgments of most people could be mistaken about the right thing to do in one or more of the scenarios, but we disagree that the Footbridge results are incorrect or exhibit a 'moral heuristic' gone awry.<sup>68</sup> To the contrary, there is a plausible principled explanation for reconciling the judgments in the two cases. Recall again that in the Standard Trolley scenario the subject is *the driver* of the trolley, not a bystander. Thus, in one scenario the subject is already involved and can either continue driving straight, killing five, or turn the trolley and kill one (most people turned). In the other scenario, the subject is a third party faced with a choice of intervening in the situation by pushing a large man to his death (most did not push) or letting the train proceed. Judith Jarvis Thomson argues for a principled moral distinction between the driver (permissible to turn the trolley), on the one hand, and the footbridge and bystander cases (both not permissible), on the other.<sup>69</sup> According to Thomson, the 'killing versus letting die' principle justifies the difference.<sup>70</sup> The driver must kill one rather than kill five, but a bystander who could push a large man or pull a lever to turn the train must let the five people die rather than kill one person. Thus, it is at least plausible that the results from the two

<sup>67</sup> Berker, *supra* note 44, at 329.

<sup>68</sup> To conclude that the Footbridge scenario is an example of a 'moral heuristic' failure requires some prior, non-question-begging argument about what the correct result ought to be. We are aware of no such arguments demonstrating that the correct judgment in each case is the utilitarian one. As a general matter, we do not think an appeal to intuitions, or the brain areas correlated with intuitions, is the way to evaluate normative conclusions regarding any of these issues. See HERMAN CAPPELEN, *PHILOSOPHY WITHOUT INTUITIONS* 158–63 (2012) (explaining why the primary philosophical arguments about the trolley problems do not appeal to intuitions).

<sup>69</sup> Judith Jarvis Thomson, *Turning the Trolley*, 36 *PHIL. & PUB. AFFAIRS* 359 (2008). As an empirical matter, however, most test subjects also pull the lever in the bystander case. For the results of several experiments and variations of the trolley problem, see MIKHAIL, *supra* note 23, at 319–60.

<sup>70</sup> Thomson, *supra* note 69, at 367.

scenarios Greene et al. tested do not reveal any problems that the law must be concerned with resolving.

If Thomson's argument is sound, this may suggest other potential problems. Most significantly, subjects in the bystander case do often judge that pulling the lever is permissible,<sup>71</sup> and they may be mistaken to do so.<sup>72</sup> But notice that if this is so (and whether it is so is beyond the scope of our discussion<sup>73</sup>), then the mistake subjects are making is that they are being too utilitarian and not deontological enough—exactly the opposite normative conclusion that some legal scholars draw from the Greene studies. This again raises our fundamental point: how the law ought to respond to conflicting moral judgments depends on philosophical arguments, not on empirical information about the brain. Whether the reasoning in these arguments is valid or justified depends on the relations of propositions, not the firing of neurons. Evidence about the brain may sometimes be relevant to moral issues, but it is a conceptual mistake to presuppose that the moral questions that face the law can be answered by looking in the brain.

### III. Neuroeconomics

Similar to moral decision making, a neuroscientific literature focuses on economic decision making,<sup>74</sup> and scholars have likewise attempted to apply neuroeconomic insights to law.<sup>75</sup> The field investigates the neurological activity of people while they are

<sup>71</sup> See MIKHAIL, *supra* note 23, at 319–60.

<sup>72</sup> Thomson argues that turning the trolley is impermissible in the bystander cases because the subject is choosing to make the one person killed pay a cost the bystander would likely not himself be willing to pay. Thomson, *supra* note 69, at 366. She concludes it is thus no more permissible than stealing from someone else in order to give the money to charity. Unlike the bystander, who may permissibly decline to do a good deed (throwing himself or the large man in front of the train), the driver must not kill five people if she can kill one instead. Thomson speculates that the drastic means undertaken by the bystander may explain the distinction between the lever and footbridge cases.

<sup>73</sup> Our point is not to take sides on this particular issue, but rather to endorse the more general methodological point that the normative issues raised by the trolley problems are matters to be resolved by (empirically informed) philosophical arguments, not appeals to brain activity underlying intuitions and judgments. Neuro-reductionist approaches to morality also sometimes rely on a false dichotomy to support their cases. See, e.g., Goodenough & Prehn, *supra* note 31, at 83:

[T]his assertion begs a question [*sic*]: if ought is something more than the conclusion of a particular kind of natural mental process, where does that something more come from? Even the Kantian move to duty, rationalism and universals merely shifts the exercise from one mental process to another. In all of its forms, this train of argument attributes to moral standards an independence from physical causation in the discoverable processes of the brain. And the question remains: if not physical processes, then what? At heart, the romantic approach rests on an often unacknowledged spiritualism.

Neuro-reductionism and spiritualism is a false dichotomy with regard to morality. For a non-spiritual, naturalist, non-neuro-reductionist account of morality, see Philip Kitcher, *The Ethical Project* (2011).

<sup>74</sup> See Alan G. Sanfey et al., *Neuroeconomics: Cross-Currents in Research on Decision-Making*, 10 TRENDS IN COG. SCI. 108 (2006); Alan G. Sanfey et al., *The Neural Basis of Economic Decision-Making in the Ultimatum Game*, 300 SCI. 1755 (2003). See also Ariel Rubinstein, *Comment on Neuroeconomics*, 24 ECON. & PHIL. 485 (2008) ('Neuroeconomics will remain a hot topic in economics for the coming decade, probably one of the hottest').

<sup>75</sup> See Jedediah Purdy, *The Promise (and Limits) of Neuroeconomics*, 58 ALA. L. REV. 1 (2006); Morris B. Hoffman, *The Neuroeconomic Path of the Law*, in LAW & THE BRAIN (S. Zeki & O. Goodenough eds., 2006); Terrence Chorvat & Kevin McCabe, *The Brain and the Law*, in LAW & THE BRAIN; Paul Zak, *Neuroeconomics*, in LAW & THE BRAIN; Terrence Chorvat & Kevin McCabe, *Neuroeconomics and Rationality*, 80 CHI.-KENT L. REV. 1235 (2005).

engaged in economic decision making, and is similar to the field of behavioural 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.<sup>76</sup> Rather than seeking to illuminate economic behaviour with psychological explanations, however, neuroeconomists seek to explain this behaviour with explanations of brain activity.<sup>77</sup> Legal scholars then apply these explanations to legal issues, just as behavioural economists do with psychological explanations.<sup>78</sup> According to one legal scholar, 'the 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'.<sup>79</sup>

Current neuroeconomic studies use fMRI to scan the brains of people while they make simple economic decisions. One prominent series of studies, and the example we will discuss, focuses on the 'ultimatum game'.<sup>80</sup> Here is how the game is played. Two participants are told that a particular sum of money is to be divided between them; player 1 proposes a division and then player 2 will choose to accept or reject it.<sup>81</sup> In a one-shot game, the players keep the proposed division if player 2 accepts; if player 2 rejects the offer, they both walk away with nothing.<sup>82</sup> According to the classic rational-actor model of economic decision making, the rational thing for player 1 to do is to propose that player 2 should take the smallest unit above zero, and that player 1 should therefore then keep the rest for himself. This is the 'rational' thing for player 1 to do because (1) this maximizes player 1's share (and maximization is the ultimate goal according to the model), and (2) it is rational for player 2 to accept any amount offered greater than zero (because any amount will be higher than zero, and zero is player 2's only other option).<sup>83</sup> Not surprisingly, as in many other areas, people deviate routinely from the outcomes predicted by the rational-actor model. For example, in most studies about half of the player 2s who perceived an offer as unfair rejected it.<sup>84</sup>

The neuroscience research of Alan Sanfey and colleagues purports to tell us why some people engaged in the 'irrational' behaviour of rejecting offers they perceived as unfair.<sup>85</sup> The studies used fMRI scans to examine the brains of players presented with 'unfair offers'; the researchers noticed increased activity in several brain regions.<sup>86</sup> Three areas in particular that showed increased activity when presented with unfair offers were the 'bilateral interior 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

<sup>76</sup> See generally Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998).

<sup>77</sup> Sanfey et al., *Neuroeconomics*, *supra* note 74, at 108; see also Chorvat & McCabe, *Neuroeconomics*, *supra* note 75, 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').

<sup>78</sup> Chorvat & McCabe, *Neuroeconomics*, *supra* note 75; Zak, *supra* note 75; Purdy, *supra* note 75.

<sup>79</sup> Purdy, *supra* note 75, at 39-40.

<sup>80</sup> Sanfey et al., *Ultimatum*, *supra* note 74, at 1775. The ultimatum game is just one of several examples in the neuroeconomics decision-making literature. For an overview of others, see Purdy, *supra* note 75.

<sup>81</sup> Sanfey et al., *Ultimatum*, *supra* note 74, at 1775.

<sup>82</sup> *Id.*

<sup>83</sup> *Id.*

<sup>84</sup> *Id.*

<sup>85</sup> *Id.*

<sup>86</sup> *Id.* at 1756.

cingulated cortex' (which has been 'implicated in detection of cognitive conflict' such as between 'cognitive and emotional motivations').<sup>87</sup> Moreover, subjects with greater increases in these three areas were more likely to reject the unfair offers. By contrast, subjects whose brains showed increased activity in the more 'rational' areas of the brain were more likely to accept the offers.<sup>88</sup>

According to Terrence Chhorvat and Kevin McCabe, the results support a particular conception of economic (and, more generally, human) decision making as the product of different brain processes in competition with one another: emotional and rational.<sup>89</sup> This is the same distinction drawn by Greene and colleagues with regard to moral decision making. The 'key questions', Chhorvat and McCabe explain, are 'how does the brain decide which problems it will address?' and 'what neural mechanisms are used to solve the problem?'<sup>90</sup> With regard to the ultimatum game, different brain regions (the emotional and rational) 'seem to embody different thought processes'.<sup>91</sup> Furthermore, they contend, because the 'anterior cingulate cortex' (ACC) (which is 'clearly involved in cognitive conflict resolution'<sup>92</sup>) was 'significantly active' for both subjects who accepted and rejected the offers,<sup>93</sup> the ACC therefore 'seems to moderate between these different regions'.<sup>94</sup> We can summarize the neuroeconomic explanation of what occurred for each subject as follows: the subject is faced with an unfair offer, the subject's brain faces a decision of what to do, and so the subject's brain decides which process to use to decide this question. The two brain processes—the emotional and the rational—begin to analyse the offer. If the processes reach contradictory conclusions, then a third part of the brain adjudicates between them, deciding whether to accept or reject the offer. The researchers define the issues for future research as 'under what circumstances do these various systems [in the brain] cooperate or compete? When there is competition, how and where is it adjudicated?'<sup>95</sup>

What implications do these studies have for law? Consider two examples. Chhorvat and McCabe argue that the findings may help to construct legal regulations that assure greater compliance by citizens as well as greater social pressures to conform to legal norms.<sup>96</sup> According to this line of argument, greater compliance would presumably follow from regulations less likely to trigger the same kinds of emotional responses that caused people to reject offers in the ultimatum game. (These same kinds of emotional responses would presumably also generate social pressure regarding conformity to legal norms; deviations would produce the same types of emotional reactions as unfair offers in the ultimatum game.) One example concerns settlement negotiations and 'tort reform'.<sup>97</sup> Kevin McCabe and Laura Inglis argue that the neuroeconomic research is relevant to, and supports, encouraging parties to accept 'rational'

<sup>87</sup> *Id.* at 1756–7.      <sup>88</sup> *Id.* at 1757–8.

<sup>89</sup> Chhorvat & McCabe, *Neuroeconomics*, *supra* note 75. The neuroeconomics literature provides an interesting example of how the vocabulary of one discipline gets transposed into another (e.g. talk of 'competition' among brain locations).

<sup>90</sup> *Id.* at 1248.      <sup>91</sup> *Id.* at 1253.      <sup>92</sup> *Id.* at 1249.      <sup>93</sup> *Id.* at 1253.      <sup>94</sup> *Id.*

<sup>95</sup> Sanfey et al., *Neuroeconomics*, *supra* note 74, at 114.

<sup>96</sup> Chhorvat & McCabe, *The Brain*, *supra* note 75, at 127.

<sup>97</sup> Kevin McCabe & Laura Inglis, *Using Neuroeconomics Experiments to Study Tort Reform*, Mercatus Policy Series (2007), available at [http://mercatus.org/sites/default/files/20080104\\_Tort\\_Final.pdf](http://mercatus.org/sites/default/files/20080104_Tort_Final.pdf).



settlements,<sup>98</sup> rather than allowing emotions to cause parties to reject what they perceive to be 'unfair' offers and irrationally 'leave money on the table' (like the person in an ultimatum game who refuses a \$1 offer and walks away with nothing).<sup>99</sup> A second proposed example concerns property rights. Paul Zak suggests that neuroscience may provide 'neural clues' about irrational behaviour towards property, such as why 'people pay more to protect property than the expected loss associated with its expropriation'.<sup>100</sup> In sum, one major policy concern underlying neuroeconomic discussions is that 'emotional' areas in the brain cause people to make decisions that deviate from the calculated decisions implied by rational-actor models of behaviour.

As with moral decision making, we resist drawing any normative conclusions from these studies for law. The characterizations of what the studies purport to show run into a number of conceptual problems. Some of these problems are similar to those discussed above regarding moral decision making.

First, the fact that unfair offers produced negative emotional reactions in subjects does not mean that the emotions, or the brain areas associated with emotions, caused the subjects to reject the offers. In the ultimatum-game studies, the data show what the subjects' brains were doing while they (the subjects) were deciding whether to accept or reject the offer. Consider the following analogy. Suppose a person's face turned red whenever he was angry. Now, suppose when faced with an unfair offer in the ultimatum game, his face turned red and he rejected the offer. Surely we would not say this is evidence that the person's face rejected the offer; similarly, why then conclude that a subject's insula cortex rejected the offer because there was activity in that area on a brain scan?<sup>101</sup> The emotional reactions could have merely accompanied decisions to reject offers otherwise judged to be unfair by subjects. In other words, the emotions could have been effects, not causes.<sup>102</sup>

Second, even if the emotional reactions caused subjects to reject the offers, the emotions could be based on prior judgments about the offers themselves.<sup>103</sup> People may react emotionally to what they perceive or judge to be just and unjust, fair and unfair, and these reactions are affected by a person'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 even if subjects rejected unfair offers because of their emotional reactions, the reactions may themselves have been caused by judgments about

<sup>98</sup> In conventional law-and-economics analysis, a 'rational' settlement is one that for plaintiffs exceeds the expected outcome at trial (plus costs), and that for defendants is one that is less than this expected amount (plus costs).

<sup>99</sup> *Id.*

<sup>100</sup> Zak, *supra* note 75. See also Goodenough & Prehn, *supra* note 31, at 98–100 (suggesting that neuroscience research may illuminate normative attitudes about intellectual property).

<sup>101</sup> Similarly, if an increased heart rate occurs when someone is lying, we would not (for good reason) say his heart is causing him to lie.

<sup>102</sup> See Poldrack, *supra* note 59 (discussing limitations of drawing 'reverse inferences').

<sup>103</sup> Emotions have objects and causes. These may, but need not, be the same thing. For example, a loud noise may be the *cause* of one's fear, but the *object* of that fear may be the possibility that there is a burglar in the house. See M.R. BENNETT & P.M.S. HACKER, *PHILOSOPHICAL FOUNDATIONS OF NEUROSCIENCE* 206 (2003) ('what 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').

the unfair offers.<sup>104</sup> Having a properly working brain (including a properly working insula cortex) may make it *possible* for one to have this emotional reaction, but the reaction may be a link in a causal chain connecting a moral judgment and behaviour.

Third, as with the fMRI studies on moral judgments, it begs the question to assume that the 'emotional' judgments are incorrect. It also begs the question to presuppose the law ought to limit such judgments and foster the 'rational' judgments associated with 'cognitive' brain processes.<sup>105</sup> The complex normative questions underlying legal theory and policy cannot be sidestepped with an appeal to 'competition' among brain processes and the need to make more 'rational' decisions.

There are deeper conceptual problems with some of the characterizations of the neuroscientific results. The descriptions of two competing brain processes, with a third area 'adjudicating' conflicts, are examples of what Max Bennett and Peter Hacker refer to as the 'mereological fallacy'.<sup>106</sup> The fallacy arises when attributes that are ascribed to a part of a person make sense only when ascribed to the person as a whole.<sup>107</sup> It makes no sense to say that a brain or a brain area 'decides', 'reasons', or 'adjudicates'. We know what it means for a *person* to make decisions, to consider reasons, and to adjudicate disputes, and we know that the person needs a brain to do these things. But we do not know what it means for the anterior cingulate cortex to decide, reason, or adjudicate because no sense has been given to such claims.<sup>108</sup> Until sense has been given to what it means for an area of the brain to 'adjudicate' a conflict—and this meaning will differ from what we ordinarily mean by 'adjudicating' and so license different inferences—then an empirical investigation of where 'competition' in the brain is 'adjudicated' is bound to fail.<sup>109</sup> Claims must make sense before they can be true or false. The current neuroeconomic explanation of decision making misguidedly ascribes psychological attributes to the brain (e.g. deciding, reasoning, adjudicating) that only make sense when attributed to the person. This confusion undermines attempts to draw conclusions for law.

In addition to ascribing human behaviour to parts of the brain, neuroeconomic explanations also go a step further and ascribe behaviour of groups of people to areas in the brain. Consider the following description from Sanfey and colleagues:

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

<sup>104</sup> *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 it. And one knows it to be unjust because it rides roughshod over someone's rights, not because one flushes in anger.')

<sup>105</sup> This question-begging feature is not unique to neuroeconomics. Reliance on a narrow conception of 'rationality' in normative arguments is a common feature of economics-inspired legal scholarship. See Michael S. Pardo, *Rationality*, 64 ALA. L. REV. 142 (2012).

<sup>106</sup> *Id.* at 133–4.      <sup>107</sup> *Id.*

<sup>108</sup> Characterizing a brain area as engaging in this behaviour leads to conceptual confusion. We know, for example, that a person needs a brain to adjudicate disputes. Does the anterior cingulate cortex also need its own brain to adjudicate (which would have its own ACC, which adjudicates with ... ad infinitum)?

<sup>109</sup> Compare an experiment to determine whether a Dworkinian principle 'weighs' more than an elephant, or a judge's assertion that she will decide the case 'in her brain'. Until sense has been given to 'weight', 'in the brain', 'adjudicates', or 'competition' in the various claims, we cannot empirically investigate whether what is expressed (or what is trying to be expressed) is true or false.

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.<sup>110</sup>

For the same reasons that 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. Corporations, and other groups of people, act through the intentional behaviour of individuals. The analogy of the brain to the modern corporation is more confusing than illuminating—despite the superficial similarities, no part of the brain acts with the type of intentionality that explains the behaviour of human actors in a modern corporation. Ascribing human behaviour to brain activity brings conceptual confusion rather than empirical elucidation to legal theory.

### Conclusion

In this chapter, we have examined claims for the importance of neuroscience for law in three distinct areas: jurisprudence, morality, and economic decision making. In each realm, we have drawn the conclusion that the arguments being evaluated involve examples of overclaiming.<sup>111</sup> We are not sceptical about neuroscience as a discipline nor do we deny that, in time, neuroscience may play a greater role in law and public policy. As we have done in other work,<sup>112</sup> we caution against the sorts of overclaiming that are a frequent feature of the literature. Our motivation here, as always, is to improve the quality of the arguments made for and against an increased role for neuroscience in law.

<sup>110</sup> Sanfey et al., *Neuroeconomics*, *supra* note 74, at 109.

<sup>112</sup> See Pardo & Patterson, *supra* note \*.

<sup>111</sup> See Morse, *supra* note 1.