Law and Neuroscience: Progress, Promise, and Pitfalls

Owen D. Jones¹,³ & Anthony D. Wagner²,³

A chapter forthcoming in:
THE COGNITIVE NEUROSCIENCES, 6TH EDITION
Gazzaniga, Mangun, and Poeppel, eds
MIT Press, 2019

[version of July 6, 2018]

Abstract:

This chapter provides an overview of new developments at the interface of law and neuroscience. It describes what is happening, explains the promise and potential influences of neuroscientific evidence, and explores the contexts in which neuroscience can be useful to law. Along the way, it considers some of the legal problems on which neuroscientific data are thought, at least by some, to provide potential answers, and highlights some illustrative cases. It also surveys emerging research that documents how interdisciplinary teams of legal scholars, judges, and neuroscientists are yielding progress and identifying potential pitfalls.

¹New York Alumni Chancellor’s Chair in Law & Professor of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37203; ²Professor, Department of Psychology & Stanford Neurosciences Institute, Stanford University, Stanford, California 94305; ³MacArthur Foundation Research Network on Law and Neuroscience

Correspondence: ODJ, owen.jones@vanderbilt.edu; ADW, awagner@stanford.edu

Acknowledgements: This work was supported in part by a grant to ODJ from the John D. and Catherine T. MacArthur Foundation, and a gift from the Glenn M. Weaver Foundation to Vanderbilt Law School. Its contents do not necessarily represent the official views of the MacArthur Foundation or the MacArthur Foundation Research Network on Law and Neuroscience or the Weaver Foundation. We are grateful to Peter Imrey for helpful comments, and to Emily M. Lamm for research assistance.
Introduction

Cognitive neuroscientific discoveries about mind and brain not only advance scientific theory, but also hold promise to inform, and often directly bear on, real-world problems of the human condition. This is increasingly evident at the intersection of law and neuroscience. The law often concerns itself with making judgments about human behavior, and the cognitive neurosciences aim to explain the psychological and neurobiological mechanisms that give rise to thought and action. The legal system — including legal decision makers (such as judges and juries) and legal policy makers (such as legislators) — is frequently charged with making decisions based on limited or noisy evidence. Given the challenges of doing so, the hope has naturally arisen that cognitive neuroscientific advances may yield informative evidence that facilitates fact-based legal decisions and policy.

While neuroscientific evidence, such as the presence of a neural injury or disorder, has long been a staple of tort law (the law of injuries), the remarkable neuroscientific advances made in recent decades have not gone unnoticed by the legal community. Increasingly, legal actors are offering neuroscientific evidence during litigation and citing neuroscientific studies during policy discussions. It appears that such evidence often has some influence on outcomes. In a complementary manner, cognitive neuroscientists are coming to appreciate how their approach can be leveraged to address important problems the law regularly confronts, as well as how their methods and results may be used, for better or worse, by legal actors.

In this review, we provide a high-level summary of recent activities at the interface of law and neuroscience, including overviews of what is happening, of the potential influences of neuroscientific evidence, and of contexts in which neuroscience can be useful to law. Along the way, we consider some of the legal problems on which neuroscientific data are thought, at least by some, to provide potential answers, and we highlight some illustrative cases. Throughout, the chapter reflects our view that there is a zone of suitable sense that lies somewhere between being too zealous about the long-term effects of neuroscience on law and being too skeptical that neuroscience has anything useful to offer.

Cross-field Interactions — The Emergence of “Neurolaw”

We begin by considering some of the key developments that have propelled interactions between neuroscience and law over the last ten to fifteen years.

First, as already noted, lawyers are increasingly offering neuroscientific evidence in the courtroom. In the civil (non-criminal) domain, for example, one core issue of the multi-district NFL concussion litigation concerns the neurological effects of repetitive impacts to the head (In re: NFL Players’ Concussion Injury Litigation, 2016; Grey & Marchant 2015). Neuroscience also appears in contexts as varied as medical malpractice litigation, on one hand, and suits seeking disability benefits, on the other. In the criminal domain, many defendants now offer evidence of brain abnormalities — such as tumors, cysts, or unusual features — to argue during the sentencing phase of a trial that they should receive a lesser punishment than would someone who acted identically, but with a “normal” brain. Former mayor of San Diego Maureen O’Connor, for instance, claimed that a tumor contributed to her gambling addiction, which in
turn led to the embezzlements of which she was convicted (U.S. v. Maureen O’Connor). The past decade has even seen attempts to enter functional brain imaging evidence purported to reveal the veracity of a defendant’s testimony, a development to which we return below.

In 2007, the John D. and Catherine T. MacArthur Foundation funded the interdisciplinary Law and Neuroscience Project (under Michael Gazzaniga, and later Owen D. Jones, Directors) to help build direct links between neuroscience, psychological science, academic law, and legal actors such as judges and attorneys. In 2011, the foundation funded the new Research Network on Law and Neuroscience (Owen D. Jones, Director). Over 12 years, these efforts propelled exploration of the promise and the limitations of using neuroscientific research to further the goals of criminal justice, building bridges between neuroscientists and legal scholars. Together with leading federal and state judges, teams co-designed and published dozens of legally relevant experiments, as well as many analyses and proposals for ways the legal system could use neuroscience usefully, while simultaneously minimizing misuses. (See www.lawneuro.org for details, including members, publications, resources, and more.)

Given rapid expansion in the types and technical complexity of neuroscientific evidence available, along with the growth in their submission as evidence, cross-field education is critical. Some of this, of course, will come in the form of expert witnesses, when neuroscientists share knowledge with the legal system, in the context of specific litigation (Jones et al, 2013a). But more broadly, this education often takes the form of training sessions and seminars. For example, a number of organizations have offered, and judges are increasingly requesting, some basic exposure for judges in the technologies, vocabularies, capabilities, and limitations of neuroscientific techniques. Over the past decade, more than 1,000 judges — along with many legal scholars, prosecutors, and defense attorneys — have participated in training sessions offered by the American Association for the Advancement of Science, the Federal Judicial Center, the MacArthur Foundation Research Network on Law and Neuroscience, and the MacArthur Law and Neuroscience Project.

Finally, burgeoning activity in law and neuroscience (sometimes called “neurolaw”) is evident along other critical dimensions. To give but a few examples, for context, consider that neurolaw publications numbered barely 100 in 2005, but swelled sixteen-fold over the next decade, to over 1600 today. Across the same time span, over 150 law and neuroscience conferences and symposia were hosted, a variety of law and neuroscience societies formed around the globe, and a number of law schools and other departments started offering neurolaw courses, some using a dedicated textbook on the subject (Jones, Schall, and Shen, 2014a). Broader knowledge sharing has taken the forms, for instance, of cover-page articles in The New York Times Sunday Magazine (Rosen 2007) and the American Bar Association Journal (Davis, 2012), a multi-part television program, various radio documentaries and interviews, a complimentary electronic newsletter (Neurolaw News) and more than 50 neurolaw video lectures (at https://www.youtube.com/user/lawneuroorg).

**Driving the Interest**

There are doubtless many drivers of the increased interest in neurolaw. But at the most basic level, it arises from the intersection of: a) perennial questions that the legal system has been
grappling with for generations; and b) the proliferation of new neuro-technological capabilities. Where these overlap springs the hope — or, at the very least, active curiosity — that neuroscientific tools that can be applied to humans may yield better answers to some legally relevant questions that have historically yielded unsatisfying or uncertain solutions.

For instance: Is this person responsible for his or her behavior? What was this person’s likely mental state at the time of the act? How competent is this person? Is this person lying? What does this person remember? How accurate is this person’s memory? Is this person really in pain, and — if so — how much? How can we improve juror and judge decisions?

And what developments have laid foundation for the hope that cognitive neuroscience can help answer these questions? For one thing, many people — including legal thinkers — increasingly recognizes that the brain is not a product of either nature or nurture, but rather necessarily exists at the intersection of genes and environments. They increasingly understand that the brain is the product of evolutionary processes, including natural selection, that have shaped it to readily associate various environmental inputs with behavioral outputs that tended (on average, in past environments) to increase the chances of survival and reproduction. And they increasingly understand that human cognition and behavior — including both relatively “automatic”, non-conscious phenomena (e.g., implicit racial biases) and more “controlled”, conscious phenomena (e.g., planning future acts) — are products of the brain, with some emerging from functionally specialized neural processes and others from large-scale network computations.

Against this background, there also has been increasing awareness of the remarkable rate of technological progress in the neurosciences. This includes awareness of key new tools of cognitive neuroscience that provide unprecedented insights into how human minds and brains work, as well as unique opportunities to try to ‘read out’ from neural signals what a person is perceiving, thinking, or remembering (e.g., Norman, Polyn, Detre, & Haxby, 2006; Naselaris, Kay, Nishimoto, & Gallant, 2011). These cutting-edge tools — including brain imaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), and data analytic methods such as machine learning, as well as the combination of both kinds of methods — have yielded both striking new discoveries as well as overhyped illusory advances. In turn, cognitive neuroscience’s many discoveries and advances have, for better or for worse, tantalized the legal system with the prospects of answering some of its most challenging questions, and commensurate concerns for associate risks (Aharoni, Funk, Sinnott-Armstrong & Gazzaniga, 2008; Gazzaniga 2008; Greeley 2009; 2013; Brown & Murphy 2010; Freeman 2010; Goodenough & Tucker, 2010; Seki & Goodenough; Blitz 2010; 2017; Farahany, 2011; Moore 2011; Morse, 2011; 2013; Morse & Roskies 2013b; Denno 2015; Patterson & Pardo 2016; Slobogin 2017; Alces 2018).

Illustrative Research

In this section we provide a sampling, for general flavor, of some of legal problems on which neurolaw experiments have been published in the last decade or so. We focus on the works with which we are most familiar, given that we each served on the MacArthur Foundation Research Network on Law and Neuroscience (the “Network”). Readers interested in the broader neurolaw
Brain-based Memory Detection

Behavioral expressions of memory serve as critical evidence for the law, including eyewitness identifications and memory-based statements about an individual’s intent or frame-of-mind during a past act (National Research Council, 2014). Mnemonic evidence is often challenged by the opposing side, leaving the jury to decide whether to believe, and how heavily to weigh, the evidence. Given this long-standing challenge for the law, there is interest in whether neural measures can detect the presence or absence of a memory, or distinguish true from false memories (Nadel & Sinnott-Armstrong 2012; Lacy & Stark, 2013; Schacter & Loftus, 2013). Being able to detect reliable neural signals of memory could be useful in a variety of investigative contexts, including probing the probability of deception (see next subsection).

To examine whether functional brain imaging can be used to detect real-world memories, one Network working group, led by one of us (Wagner), put cameras around the necks of undergraduate students that automatically took photos as they navigated their lives for a few weeks (Rissman et al, 2016; see related work by St. Jacques et al., 2011; St. Jacques & Schacter, 2013). Subsequently, selected photos from a subject’s camera were interleaved with photos from other subjects’ cameras and displayed while the subject made memory decisions during fMRI. Machine-learning techniques applied to the fMRI data — here, multivoxel pattern analyses — revealed that activity patterns in numerous cortical regions along with the medial temporal lobe can be used to classify whether the subject is viewing and recognizing photos of their past (i.e., hits) versus viewing and perceiving as new photos from someone else’s camera (i.e., correct rejections). Classifier accuracy was well above chance (approaching ceiling performance in some cases) and, intriguingly, this was the case even when the classifier was applied to brain data from subjects other than the ones on which it was trained up. In addition to detecting memories for real-world autobiographical events, a lab-based study revealed high accuracy when classifying brain patterns associated with recognizing studied faces versus correctly rejecting novel faces, as well as discriminating higher confidence versus lower confidence memories (Rissman et al, 2010).

While the above findings suggest that, under controlled experimental conditions, memory states can be detected from fMRI-measured brain patterns, initial studies also point to important boundary conditions. First, while high classification accuracy is possible (under some conditions) when discriminating recognized stimuli from stimuli perceived as novel, classification accuracy was only slightly above chance when attempting to discriminate true versus false recognition of faces (Rissman et al, 2010). This finding converges with a wealth of other data highlighting the similarity of brain responses during true and false memory (Schacter & Slotnick, 2004), and suggests that brain-based measures may not solve the law’s frequent quandary of knowing when a witness’s memory is accurate or mistaken. Second, classification accuracy was essentially at chance when applied to implicit memory — i.e., discriminating between old stimuli that a subject failed to recognize (i.e., misses) from new stimuli perceived as novel (i.e., correct rejections) (Rissman et al, 2010). Finally, the high level of fMRI-based classification of hits versus correct rejections fell to chance when subjects used cognitive
countermeasures (shifting how they attended to memory) in an effort to mask their neural patterns of memory (Uncapher et al., 2015). As with the polygraph (National Research Council, 2003) and fMRI-based lie detection (see below), potential real-world application of brain-based memory detection can be ‘beat’ by motivated non-compliant individuals. Thus, while extant data highlight that brain-based memory detection is possible, significant hurdles to real-world application remain.

**Brain-based Lie Detection**

As noted at the outset, lawyers are increasingly proffering (i.e., “offering into evidence”) neuroscientific evidence, both structural and functional. In many cases, such evidence is the subject of admissibility hearings, in which a judge determines (according to state or federal law standards) whether the jury will be allowed to hear and see the evidence. For instance, in the case of *United States v. Semrau* (2010), the defendant Lorne Semrau, who ran a psychiatric group, was prosecuted for Medicare and Medicaid fraud. Although not all criminal statutes require knowledge of wrongdoing to be guilty, it is in fact one element of proving fraud that Dr. Semrau have known that what he was doing was illegal. In his defense, Dr. Semrau sought to introduce a report from the company Cephos purporting to show that an fMRI lie-detection protocol “indicated he is telling the truth in regards to not cheating or defrauding the government.” Following 16 hours of hearings before a magistrate judge, the magistrate convincingly recommended to the trial judge that the evidence be excluded from the jury, due to specific flaws in the particular protocol, as well as doubt that the urged inferences could properly be drawn from the results (Shen & Jones, 2011).

With the advent of fMRI, cognitive neuroscientists are examining whether brain-based lie detection is possible. Despite some very promising studies (Greene & Paxton, 2009), the prospects for legal use remain almost entirely speculative (Bizzi et al., 2009; Wagner, 2010; Wagner et al., 2016). Take-home points from the literature (Christ et al., 2009; Farah et al., 2014) include: (a) laboratory-based studies predominantly use instructed or permitted lie paradigms, and have negligible stakes for failure to successfully deceive (in contrast to the stakes in real-world settings); (b) a set of frontal and parietal lobe regions are often more active during the putative “lie” versus “truth” conditions, and most evidence comes from group-based analyses that average over trials and subjects (c.f., the law requires an assessment of truthfulness about individual facts in individual brains); (c) experimental design limitations raise uncertainty as to whether these neural effects reflect responses associated with deception or whether they reflect attention and memory confounds that are unrelated to deception; and (d) countermeasures appear to alter these neural responses, suggesting that, even if associated with deception, it may be possible to mask such responses. These limitations will frequently prevent brain-based techniques from satisfying the legal standards for admissibility of scientific findings. Indeed, some of these limitations and boundary conditions, along with others, were considered in the *Semrau* case, as well as the handful of other cases in which judges decided not to admit fMRI-based “lie detection” testimony into evidence.
Detection and Classification of Mental States

Generally speaking, the government must prove, in order to get a criminal conviction, both that a defendant performed a prohibited act (“actus reus”) and that he did so in one of several defined states of mind (“mens rea”) (for more on this, see Morse & Newsome, 2013). Because most crimes are matters of state law rather than federal law, the mental state definitions can vary. However, the “Model Penal Code” — which itself has no legal force — has been widely influential on the mental state definitions in most states. By its taxonomy, culpable mental states include: purposeful, knowing, reckless, and negligent — in descending sequence of severity, each with importantly different sentencing results. In Colorado, for instance, the difference between being convicted of a knowing homicide on one hand or a reckless homicide on the other could mean the difference between 14 years in prison and incarceration-free probation.

Scholars have long debated whether the knowing-versus-reckless distinction drawn by law actually exists in the brains of defendants, a concern heightened by recent behavioral work strongly suggesting that juror-like subjects have a difficult time distinguishing between the two (Shen et al, 2011; Ginther et al, 2014; Ginther et al, 2018). Consequently, another line of research seeks to explore the extent to which coupling fMRI with machine-learning algorithms could shed light on whether there is a real psychological distinction between a “knowing” frame of mind and a “reckless” frame of mind. And one Network working group, led by Gideon Yaffe, found that the combination of fMRI and machine-learning algorithms could (under laboratory conditions) predict with high accuracy whether a subject was in knowing versus in reckless frames of mind. This arguably suggests that the distinction the law had posited academically actually exists neurologically. And this is the first proof of concept that it is possible to read out a law-relevant mental state of a subject, in a scanner, in real time (Vilares et al, 2017).

Intent and Punishment

Humans are notoriously prone to various kinds of psychological biases. At the same time, few things are more crucial to the fair administration of criminal justice than trying to ensure that jurors and judges are minimally biased in their decisions about whether or not a defendant is criminally liable (typically a decision for the jury) and, if he is, how much to punish him (typically a decision for the judge). Until recently, nothing was known about how human brains make these important decisions.

Consequently, one line of research explores the extent to which fMRI might illuminate the neural processes underlying these determinations, which could potentially be an important first step in learning how to debias them (through, for instance, more effective training interventions). A first fMRI study found correlations between guilt and punishment decisions and activity in regions commonly associated with analytic, emotional, and theory-of-mind processes (Buckholtz et al, 2008). A subsequent study suggested that theory-of-mind circuitry may either gate or suppress affective neural responses, tempering the effect of emotion on punishment levels when, for instance, a perpetrator’s culpability was very low, at the same time the harm he caused was very high (Treadway et al, 2014). A third study, using repetitive transcranial magnetic stimulation (rTMS) to test the causal role of right dorsolateral prefrontal cortex, found, as predicted, that compared to sham stimulation rTMS changed the amount that subjects punished protagonists in
scenarios, without altering how much they blamed those protagonists (Buckholtz et al, 2015). Breaking liability and punishment decisions down into constituent steps, a Network working group led by Owen Jones recently identified distinct neural responses that separately correlate with four key components of liability/punishment decisions: 1) assessing harms; 2) discerning mental states in others; 3) integrating those two pieces of information; and 4) choosing punishment amounts (Ginther et al, 2016).

Adolescent and Young Adult Brains

A constant challenge for legal systems is figuring how best to handle young offenders. While it has always been obvious that the very young are not as culpable for bad behavior as are the mature, legal systems have often struggled to develop juvenile justice regimes that are stable and fair. Several U.S. Supreme Court cases reflect this struggle. In Roper v. Simmons (2005) the Court held unconstitutional any sentence to death for a crime committed by an adolescent of 16 or 17 years old. In Graham v. Florida (2010), the Court similarly held it unconstitutional to sentence any juvenile offender, in a non-homicide crime, to a sentence of life imprisonment without possibility of parole. In Miller v. Alabama (2012) the Court went further. It held that mandatory life imprisonment without possibility of parole, for those under the age of 18 at the time of their crimes, was unconstitutional — even in cases of homicide. (However, the Court left open the possibility of such a sentence, if the judge were to make an individualized assessment of the particular juvenile, crime, and surrounding circumstances.) Although the role neuroscientific arguments actually played in the disposition of these cases is debatable (Morse, 2013), it is notable in itself that neuroscientific arguments about adolescent brain development were provided to the Court in each case, and cited in some of them (Bonnie & Scott, 2013).

Complementing structural data that suggest that full maturation of the human brain may occur as late as into one’s 20s (Gogtay et al, 2004; Mills et al, 2014), a wealth of behavioral and functional neural data highlight the context-dependence of developmental trajectories (Luna, 2012; Albert, Chein, & Steinberg, 2013). Importantly, these studies of adolescents and young adults might illuminate issues potentially relevant to juvenile and young adult justice. For example, potentially bearing on the legal system’s challenge of deciding when and how to hold juveniles criminally responsible for their behavior, a Network working group led by B.J. Casey is exploring whether it is possible to draw meaningful lines between juveniles and young adults using fMRI and behavioral assays (Casey et al, 2017). In one study (Cohen et al, 2016), fMRI data and behavioral measures from 250 juveniles and young adults examined cognitive control under affectively arousing versus neutral conditions. Among the findings was this one: the brains and behaviors of 18-to-21 year olds operate more like older adults under some environmental circumstances — specifically, when arousal and affective states are neutral — and more like juveniles in others — when arousal and affect are elevated (such as when emotion is triggered by stimuli or when performance is under peer observation). These data may have broad implications for the law, as they suggest that the age at which mature behavior may be fully realized is context dependent.

Categories of Relevance

There are at least seven contexts in which neuroscience can be relevant to law (Jones, 2013).
**Buttressing**

Neuroscientific evidence, most commonly perhaps, can be used to buttress other – typically behavioral – evidence. For example, suppose a criminal defendant has raised an insanity defense. If there is behavioral evidence consistent with insanity, those data will be the most salient evidence. If it turns out that there is also evidence of an acute abnormality in brain form or function, then the latter will buttress the former. But note that the neuroscientific evidence, no matter how strong, would be insufficient on its own to build a credible insanity defense, if there were no behavioral evidence consistent with insanity to accompany it. In such a case, the buttressing effect of neuroscientific evidence would add to the weight of the behavioral evidence, not independently supplant it; that is, the brain data could support a conclusion, but not drive it.

**Detecting**

One of the most potent uses of neuroscience, perhaps, is its ability to detect facts that may be legally relevant. For example, in the 1992 New York case of People v. Weinstein, Mr. Weinstein, an executive in Manhattan, came home one day, strangled his wife, and threw her out of the couple’s 12th story apartment building. After arrest Mr. Weinstein complained of headaches, which led to a discovery, through positron emission tomography, of a very large subarachnoid cyst, compressing his prefrontal cortex, known to be important for impulse control and executive function (Davis, 2017). Although it is unknown — perhaps unknowable — how much the cyst contributed to the murder, the possession by the defense of a visually powerful brain image contributed to Mr. Weinstein’s plea agreement with the state. And it illustrates the extent to which neuroscientific methods for detecting brain structures and functions may uncover new legally relevant avenues to pursue. The same is true, for instance, of the extent to which brain imaging might more clearly detect injuries — or even the existence and amount of pain — in torts cases (Kolber, 2007; Pustilnik 2012, 2015; Davis 2017). Of course, as noted earlier, some maintain the hope that functional neuroimaging may one day enable the reliable detection of lies or legally relevant memories.

**Sorting**

Neuroscience might also aid the legal system in sorting individuals into different categories, for different purposes. A paradigmatic example, perhaps, would be if neuroscientific measures could reliably identify criminal addicts who were most susceptible to rehabilitative interventions. In theory, the legal system could then send some such individuals into drug rehabilitation, instead of into the general prison populations.

**Predicting**

Over time, neuroscience may make important contributions to law’s efforts to predict various kinds of behaviors. For instance, two papers (Aharoni et al, 2013, 2014) provided initial evidence that certain brain-based variations in incarcerated individuals predict some of the variance in the probability of their rearrests after release. It was a small part of the variance, and the magnitude of the effect is debated due to questions about analytic approach (Poldrack, 2013; Poldrack et al., 2013).
2017). Nevertheless, as parole boards, for instance, sometimes expand and revise their actuarial approaches to predicting recidivism (including age, sex, type of crime, etc.), such observations raise the possibility that at some point in the future neuroscientific measures may become relevant. Determination of if and when such application emerges will be informed by meaningful debates about how best to interpret and apply neuroprediction (Nadelhoffer et al 2010; Poldrack, 2013; Singh et al 2013; Slobogin 2013; Poldrack et al, 2017).

**Intervening**

In theory, neuroscience could aid law through the development and validation of intervention approaches. For example, if certain drug treatments prove to substantially decrease the probability of recidivism, psychopharmacological interventions may be recommended for inclusion as a condition of parole.

Of course this, like many aspects of neurolaw, can raise important ethical considerations about what trade-offs we as a society are willing to make, between perceived benefits, attendant risks and costs, and individual rights (Illes, 2017; Morse, 2017).

**Explaining**

Neuroscientific methods are beginning to uncover regions of the brain, neural responses, and interactions within and between regions that subserve the processes by which decisions — key to the functioning of law — are made (Heekeren et al, 2008; Shadlen & Kiani, 2013). As discussed above when considering adolescent brain development, these could provide new insights into why and how individuals transgress the law, in criminal or civil domains (Scott & Steinberg, 2008; Scott et al, 2016, Steinberg 2016). Such discoveries also could provide new insights into the experiences of individuals who have been wronged. And, as noted above, they could provide insights into the processes by which jurors and judges make their decisions. All of these might increase the knowledge base on which new behavioral interventions and legal policy are deployed in furtherance of improving decisions, and the legal consequences they create.

**Challenging Assumptions in the Legal System**

Neuroscience may sometimes challenge assumptions in the legal system. For example, the legal system currently assumes that solitary confinement is insufficiently damaging to the brain to constitute “cruel and unusual punishment,” and thus it is not prohibited as unconstitutional. Perhaps that’s right. Perhaps it isn’t. The tools of neuroscience may eventually help us to know which. If the assumption is wrong, that may provide impetus for law reform.

Similarly, note that the rules of evidence can be thought of as designed to keep certain information from entering the brains of jurors, because of assumptions about how that information might affect the decisions of jurors. The evidentiary rules also reflect underlying neuroscientific assumptions about witness brains. For instance, a general rule of evidence, known as the prohibition against “hearsay,” typically operates to prevent person A from testifying as to what person B said they observed, at the time of an act relevant to the trial (such as the name of a perpetrator). The logic is that (so long as person B is available), person B’s
testimony is deemed to be more reliable than person A’s. But there are some exceptions. Among them is the exception for “excited utterances.” That exception allows person A to testify as to what person B said — so long as person B was excited, and believed to be more or less blurting things out at the time. The explicit assumption underlying this rule is that person B, being in an excited state, would not have time to lie about what she was witnessing. Perhaps that’s true. Perhaps it isn’t. The tools of cognitive neuroscience might help us to know which. And if the assumption is wrong — with respect to this evidentiary rule or others — neuroscience may again provide potential foundation for reform.

Two Key Caveats

There are, of course, many cautions and caveats about whether neuroscientific information should directly impact legal decisions and policy and, if so, how to carefully, sensibly, and responsibly incorporate such information (Jones et al, 2009; Campbell, 2013; Morse 2013; Faigman et al, 2014). For example, above we described some of the open questions and potential boundary conditions surrounding brain-based memory and lie detection. In each of the areas of research we briefly considered, as well as others being explored in by the field, additional cautions and caveats are warranted.

Here we consider two especially salient, cross-cutting caveats.

The Long Chain of Inference

First, it is not a simple thing to reason from the presence of a brain feature (a large subarachnoid cyst, for example) to the conclusion that that feature contributed meaningfully to generating or enabling a specific behavior (such as murder). Such a conclusion requires a long chain of inferences, with many potential weak links. What exactly is the brain feature at issue? How long was it there? What is known to correlate with the presence of the brain feature? What are the known causal pathways of influence? In many instances, answers to one or more of these critical questions are unknown, which greatly tempers confidence in any inferences drawn.

Unknown Frequencies of Predictors and Outcomes

Second, and relatedly, one key limitation to drawing logical and informed inferences is that the relative frequency of a feature in the population — Mr. Weinstein’s cyst, for instance — is often not known. Without that information, we have no idea how many people are walking around in the population with the same feature, but without engaging in the same behavior as did the accused. Knowing the relative frequency of a predictor, as well as of the frequency of a particular outcome (i.e., the base rate), are necessary to determine the increased likelihood, if any, of engagement in an undesirable behavior given the feature in question (National Research Council, 2003). Without this information, proper inferences are difficult to draw. With what confidence could one say that Mr. Weinstein’s cyst meaningfully, and legally, caused him to commit murder?

The issue of unknown predictor frequencies is particularly relevant given the remarkable pace of progress in neuroscientific methods in recent years. Whereas structural imaging of the human
brain has been available for a few decades, and detection of a structural abnormality is often relatively straightforward for neuroradiologists, functional imaging is a more recent development and machine-learning characterization of functional patterns and their relation to cognition is at an even earlier stage. Thus, whereas some limited information is available on the relative frequencies of structural abnormalities and their relationships to altered behavior, cognitive neuroscience is only just beginning to conduct large-scale individual difference studies of the relationships between functional brain patterns (which themselves vary depending on the particulars of the analytic approach) and cognitive states and behaviors. Early work is focused on characterizing the heterogeneity evident in healthy young adults — we seem far from the point at which we can say anything about the relative frequencies of particular functional patterns in healthy individuals and their associated outcomes, let alone those of atypical patterns and states.

**Legal Impact of Neuroscience Evidence**

In instances where neuroscientific evidence is admitted, what are its impacts? We know that jurors are, at least sometimes, significantly affected by neuroscience evidence. For example, in the case of *State of Florida v. Grady Nelson* (2010), the defendant was quickly convicted of a murder, leaving the question to the jury by simple majority vote (under Florida death penalty law, as it was at the time) whether Mr. Nelson should be executed or given life in prison without parole. With Mr. Nelson’s life hanging in the balance, the defense introduced qEEG evidence (quantified electroencephalography) in support of the inference that Mr. Nelson’s brain was too abnormal to warrant his execution. By the narrowest of possible votes, the jury gave Mr. Nelson life in prison. Afterward, two jurors granted interviews indicating that the brain data had turned their prior inclinations, to vote in favor of execution, completely around.

Some members of the judiciary are increasingly invoking neuroscience in judicial opinions (Farahany, 2015), sometimes drawing colleagues into public debates over its relevance. High profile examples include the U.S. Supreme Court cases of *Graham v. Florida* and *Miller v. Alabama* (mentioned earlier). And Supreme Court Justice Sotomayor recently referred to “a major neurocognitive disorder that compromises [the defendant’s] decision-making abilities” in her dissent from the Court’s refusal to hear the appeal in *Wessinger v. Vannoy* (2018).

Of course, given the complexity of neuroscience, one natural concern is that both judges and jurors may have a hard time understanding where it is — and equally importantly is not — relevant. Relatedly, some have expressed worry that jurors may be over-awed by the pictorial nature of some brain data, and give it more weight than it is due (Weisberg et al, 2008). Two laboratory studies investigating this phenomenon found that the images themselves appear to have no particular biasing effect on subjects — above and beyond non-pictorial neuroscientific testimony — except in the case of death penalty decisions, wherein images decreased the probability of a vote for execution (Schweitzer & Saks, 2011; Saks et al, 2014).

Given the complex interactions between law and neuroscience, there is a need for reasoned consideration of the ethical and legal impacts of neuroscientific evidence (e.g., see the 2015 report of the Presidential Commission for the Study of Bioethical Issues (2015), and selected recommendations submitted to the Commission (Jones et al, 2014b)).
Conclusions

The domains of science and law have very different goals. Painted with broad brush, these are the attempt to uncover truths, on one hand, and the attempt to fairly and effectively govern the behaviors of large populations, on the other. While truths may inform governance, they don’t dictate it. Indeed most scholars (including ourselves) believe it impossible to argue directly from a description to a prescription without reference to other values. Put another way, explanation isn’t justification. And therefore we do not expect the law will or should automatically change, or refuse to change, in light of a neuroscientific finding alone.

At the same time, advances in the cognitive neurosciences effectively guarantee a future in which the law increasingly interacts with neuroscientific evidence. Even at this relatively early stage, there is a gradual but discernible shift from nearly exclusive reliance on structural brain evidence (in cases involving any brain evidence) to increasing reliance on functional neural assays. As this shift continues to develop and accelerate, there will be divergent views on whether and when particular types of neural data should be drawn upon to inform legal decisions.

In this review, we have highlighted a few illustrative legal problems on which neuroscience research is beginning to yield potentially informative data, as well as others in which the science suggests it is premature to move from the lab to the courtroom (for other overviews, see Jones et al, 2013b; Jones et al, 2014a). Concurrently, we have considered the categories of potential relevance for neuroscience evidence, along with cross-cutting caveats. The growth of neurolaw – which crucially depends on interdisciplinary interactions — has produced significant progress and suggests promise. At the same time, there is ample cause for caution, lest over-exuberance pave a path to pitfall.

REFERENCES


DAVIS, K. (Nov. 2012). Brain trials: Neuroscience is taking a stand in the courtroom. *American
Bar Association J., 98, 36–37.


In re NFL Players Concussion Injury Litig., 821 F.3d 410 (3d Cir. 2016).


Neurolaw News. THE MACARTHUR FOUNDATION RESEARCH NETWORK ON LAW & NEUROSCIENCE.


Neuron, 44(1), 149–160.


United States v. Semrau, 2010 WL 6845092 (W.D. Tenn. June 1, 2010), aff’d, 693 F.3d 510 (6th Cir. 2012).


