Chapter 16

Neurolaw: Challenges and limits

STEPHEN J. MORSE*

Law School and Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States

Abstract

This chapter canvasses the current relevance of behavioral neuroscience to the law, especially to issues of criminal responsibility and competence. It begins with an explanation of the legal doctrines at stake. I then explore the source of the often-inflated claims for the legal relevance of neuroscience. The next section discusses the scientific status of behavioral neuroscience. Then, it addresses two radical challenges to current conceptions of criminal responsibility that neuroscience allegedly poses: determinism and the death of agency. The question of the specific relevance of neuroscience to criminal law doctrine, practice, and institutions is considered next. This is followed by a discussion of how neuroscience evidence is being used in criminal cases in five different countries, including the United States. The penultimate section points to some areas warranting modest optimism. A brief conclusion suggests that neuroscience is at present of limited legal relevance, and advances in the science might alter that judgment.

Twelve years ago, I published a state-of-the-art review of the existing neurolaw literature (Morse, 2011). The primary claim was that the major problem confronting the emerging field was the question of legal relevance. Could the translation problem be solved; could the gap be closed between the purely mechanistic discourse of neuroscience and the thoroughly folk psychological concepts of the law? This chapter will address the same issue of the potential contributions of neuroscience to law, with special emphasis on criminal justice and criminal responsibility because these are the areas that have received the lion's share of neurolaw attention. The discussion will clearly generalize to other applications, however. The conclusion remains the same as 12 years ago. At present, virtually no neuroscience addresses legal questions directly and little pure neuroscience is even indirectly legally relevant. One hopes that future advances will increase the legal usefulness of neuroscience, but the notion that neuroscience will radically change the law by replacing its concepts and procedures is a chimera. At most, one can hope for incremental reforms.

The law's criteria are virtually all behavioral—acts and mental states. This is especially true in criminal law or in any other legal context in which responsibility and competence are in question. The most relevant neurosciences are therefore the behavioral neurosciences, such as cognitive, affective and social neuroscience. There have been major advances in these fields since the beginning of the present century when noninvasive functional magnetic resonance imaging (fMRI) to investigate brain function became widely available for nonclinical research. But again, these advances have scarcely improved the legal relevance of neuroscience.

Many readers of this chapter may not be lawyers, so the chapter begins with a brief explanation of the meaning of criminal responsibility that will be used throughout. It then speculates about the source of claims for the positive influence of neuroscience. The next section discusses the scientific status of behavioral neuroscience. Then it addresses two radical challenges to current conceptions of criminal responsibility that neuroscience allegedly poses: determinism and the death of agency. The question of the specific relevance of neuroscience to criminal law doctrine, practice and institutions is considered next. This is followed by a discussion of how neuroscience evidence is being used in criminal cases

*Correspondence to: Stephen J. Morse, 3501 Sansom Street, Philadelphia, PA 19103, United States. Tel: +1-215-898-2562, Fax: +1-215-273-2025, E-mail: smorse@law.upenn.edu

in five different countries, including the United States. The penultimate section points to some areas warranting modest optimism. A brief conclusion follows.

THE MEANING OF CRIMINAL RESPONSIBILITY

This chapter employs an expansive definition of criminal responsibility that is equivalent to criminal liability, criminal blameworthiness or criminal culpability. Although the discussion will focus on US law, its structure and doctrines are remarkably similar to continental criminal law codes, albeit different terminology is used.

Crimes are defined by their criteria, what lawyers term the "elements" of the crime. Clinicians can think of these elements as analogous to the criteria for psychiatric diagnosis of the various mental disorders. The most important elements for our purposes are the act requirement (often misleadingly termed the "voluntary act") and a culpable mental state, the mens rea. For example, one basic definition of murder is the intentional killing of another human being. In this example, the prohibited act is the killing conduct, such as shooting or stabbing, and the mental state element is the purpose to kill. The most basic definition of an act is an intentional bodily movement or omission performed in a reasonably integrated state of consciousness. For example, a reflex movement or movements performed in a state of divided or partial consciousness, such as sleepwalking, would not qualify as acts for criminal law purposes, even if they resulted in harm to another. The act element has constitutional status because it is unconstitutional to punish a person for a status, that is, for some attribute of the person, such as having red hair or for being an addict (as opposed to possession of or using a controlled substance, which would be an act) (Robinson v. California, 1962).

A mens rea is not required constitutionally for less serious, regulatory crimes, but it is almost certainly required for all serious crimes involving moral turpitude (Morissette v. United States, 1952). Mens rea elements, such as purpose, intent, knowledge, conscious awareness of a substantial risk of harm (recklessness) and unreasonable unawareness of a substantial risk of harm (negligence) are crucial because they indicate the agent's moral indifference to the rights and interests of others. Harming another "on purpose" is almost always more morally blameworthy than harms caused carelessly. For the most part, these elements have ordinary language meanings.

Together, the elements are known as the prima facie case and must be proven beyond a reasonable doubt by the prosecution. If they are so proven, the defendant is prima facie criminally liable. The defendant nevertheless can avoid liability by establishing an "affirmative defense" of justification or excuse, which defenses are also defined by their criteria. The Constitution permits the state to place the burden of proof for affirmative defenses on either the prosecution or the defense.

In cases of justification, otherwise prohibited conduct is considered right or at least permissible under the specific circumstances of the case. Self-defense is a classic example because the innocent victim of wrongful aggression is justified in intentionally using proportionate force to defend against the wrongful aggression. Excuses obtain if the agent is prima facie liable because the agent has done something wrong, but the agent is not responsible. Classic examples are the insanity defense and duress. In the former, the agent was irrational at the time of the crime because, for example, he did not know right from wrong as a result of severe mental disorder. In the latter, the agent is compelled by being placed in a do-it-or-else situation through no fault of his own and a person of reasonable firmness would have yielded to the threat. Note that in cases of excuse, the defendant may have intentionally engaged in the prohibited conduct, but lacked the capacity for rationality or was coerced into doing it. For example, a defendant deluded about the need to use deadly self-defense intentionally kills his imagined attacker; a defendant compelled by a threat of death unless he kills an innocent surely intends to kill the innocent party to save his own life.

In sum, criminal responsibility in this chapter means that the prima facie case has been proven and no affirmative defense has been established. Free will as philosophers understand the term is not a criterion for criminal responsibility in any sense (Morse, 2007).

THE SOURCES OF NEUROEXUBERANCE

This section speculates based on wide reading and research in the intersection between neuroscience and law that bears on normative questions. At various times, the law has considered the findings from many sciences, including economics, sociology, different types of psychology, such as behaviorism and psychodynamic psychology, psychiatry, genetics, and now neuroscience. Although there are ethical and legal subdisciplines that have arisen as a result, such as bioethics, psychiatric ethics, and mental health law, for the most part, none of these has been based on a revolutionary approach to law or ethics. They primarily use familiar legal and ethical concepts to address traditional issues that the new sciences produce. For example, genomic information about individuals might raise acute privacy or human enhancement issues, but these are traditional questions. The most revolutionary claim arising from these sciences is typically the hoary claim that determinism is incompatible with free will and responsibility. Most of the various sciences have presented themselves

as the newest proof of determinism that allegedly should upend doctrines and practices based on personal responsibility. Advocates of this approach typically favor one form or another of consequentially-based social control that is often mischaracterized as "medical" or some analogue (Menninger, 1968; Caruso, 2021). Nonetheless, none of these has engendered the type of academic and public enthusiasm (and fear) that neuroscience has produced. The supposed challenges were best summed up by an editorial warning in *The Economist*:

> *"Genetics may yet threaten privacy, kill autonomy, make society homogeneous and gut the concept of human nature. But neuroscience could do all of these things first (The Economist, 2002)."*

The question is why neuroscience has had such a large impact.

The relation of the brain to the mind and action has been at the center of philosophical and scientific attention for centuries. The "neuroscientific" approach to understanding behavior roughly dates to the case of Phineas Gage, a railroad construction foreman who suffered a severe injury to his frontal cortex in 1848, but who miraculously survived. The traditional narrative, about which there is some doubt (Macmillan, 2000), is that, prior to the accident, Gage was a model of probity and rectitude, but that after the injury he became disinhibited, and his prior ability to plan and regulate his behavior, known as executive control, deteriorated. Today, we have a better understanding of the relation of frontal cortical function to executive control, but, even then, the case was a powerful demonstration of the relation of brain structure and function to behavior. Not until the advent of noninvasive functional magnetic resonance imaging (fMRI) in the early 1990s, however, and not really until the early 2000s, when scanners (often colloquially referred to as "magnets") became more widely available, was a noninvasive technology available that could investigate the brain function of large numbers of nonclinical subjects. As a result of the increasing availability of fMRI, there is now an immense and growing literature on the relation of brain to behavior that has fueled the scientific and popular imagination. This work seems to many people, erroneously I believe, more rigorously scientific than previous sciences of behavior. As Professor Moore (2020) emphasizes, causal brain states are much spatiotemporally closer to the behaviors they cause compared to environmental or even genetic influences. Moore adds that brain states also seem sufficient to explain behavior so that they are also a complete explanation. But is it true that we are just victims of neuronal circumstances of just a pack of neurons?

The images produced (which are *not* "pictures" of the brain) can be ravishingly arresting. In a metaphor that is question-begging because it assumes a form of mind/brain reductionism that is philosophically controversial, many enthusiasts claim we can now "look under the hood" of the acting agent to discern what the driving mechanisms are. The brain is of course the necessary biological foundation for mind and action, and we are discovering neural correlates and sometimes causes of mental states and actions, but acting human beings are usually not thought to be mere mechanisms like automobiles. The thought that we are merely mechanisms is at present scientifically unjustified and some would argue it is conceptually confused (Pardo and Patterson, 2013; Gabriel, 2017), but the possibility that this is true has created great expectations.

I speculate that there are four sources of what I have termed "neuroexuberance" among philosophers, lawyers, and others. The history of law as action-guiding is overwhelmingly one of conflict and irresolution with no method to establish an obviously right answer. There is no experiment, even in principle, to indicate that humans *should* behave in one way or another. It is all contestable. Hume's caution that we cannot derive an ought from an is still regnant (Hunter and Nedelsky, 2018). As Professor Robin Feldman usefully details, law makers often believe they lack the resources to decide the difficult normative questions, so they wrongly turn to science without understanding that science can almost never resolve such disputes (Feldman, 2009). Many commentators now seem to believe that the findings of the "hard" science of neuroscience may hold the key. Even the Supreme Court of the United States fell prey to this belief when it incorrectly distinguished neuroscience from social sciences (Miller v. Alabama, 2012, n. 5). Neuroscience and other sciences are all sciences. The important distinctions are between good and bad science and between legally relevant and legally irrelevant science.

Second, many philosophers and scientists dismiss the law's implicit folk psychology as prescientific and explanatorily empty. Folk psychology is a causal theory that explains human behavior in part by mental state variables such as desires, beliefs, intentions, plans, and willings. There is dispute about how to characterize the basic mental states, but all folk psychologists agree that mental states partially explain human behavior and are the fundamental tools that we all use to explain our own behavior and the behavior of others. The law provides one set of reasons for people to act as the law requires and the crucial criminal law responsibility criteria are virtually all acts and mental states. Thus, if folk psychology has no causal efficacy and explains nothing, then the law's behavioral premises are entirely incorrect. The neuroexuberants, such as Patricia Churchland, believe that neuroscientific understanding of behavior will finally replace folk psychology and give

the law's understanding of behavior a genuine basis at last. I return to this topic below.

Third, many philosophers and lawyers are profoundly skeptical of deontology and especially of retributive justifications for state blame and punishment (e.g., Caruso, 2021). Some incorrectly think that neuroscience proves that determinism is true, which, when coupled with hard determinist metaphysics, provides the desired conclusions that no one is really responsible for any behavior and that we should replace allegedly outmoded and unjust retributively-based responsibility practices with consequences-based social control. As noted, this argument has been made previously based on other behavioral sciences, but again, neuroscience appears to be a more "real" science that at last will provide a genuine scientific basis for the claim.

Last, behavioral neuroscience is inherently interesting and fun, albeit often difficult to perform. It provides a tangible result, not just an "argument" to which some other clever philosopher or lawyer will find a damaging and perhaps even decisive riposte. It thus offers an engaging and welcome respite from the common frustrations and annoyances of normative work.

Again, the preceding is speculation, but the amount of unjustified overclaiming and exuberance that contemporary neuroscience has produced is striking and cries out for an explanation (Morse, 2013). I have no stake in my speculations and would invite readers to speculate for themselves. I doubt that anyone will rigorously investigate the question.

THE LIMITS OF BEHAVIORAL NEUROSCIENCE

Most generally, the relation of brain, mind, and action is one of the hardest problems in all science. We have no idea how the brain enables the mind, how consciousness is produced, and how action is possible (McHugh and Slavney, 1998, pp. 11–12; Adolphs, 2015, p. 175; Cobb, 2020). The brain-mind-action relation is not a mystery because it is inherently unamenable to scientific explanation, but because the problem is so difficult. For example, we would like to know the difference between a neuromuscular spasm and intentionally moving one's arm in exactly the same way. The former is a purely mechanical motion, whereas the latter is an action, but we cannot explain the difference between the two. Wittgenstein famously asked: "Let us not forget this: when 'I raise my arm,' my arm goes up. And the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?" (Wittgenstein, 1953, § 621). We know that a functioning brain is a necessary condition for having mental states and for acting. After all, if your brain is dead, you have no mental states and are not acting. Still, we do not know how mental states and action are caused. Wittgenstein's question cannot be answered yet.

Despite the astonishing advances in neuroimaging and other neuroscientific methods—especially in understanding sensory systems and memory—we do not have sophisticated causal knowledge of how the brain works generally. The scientific problems are fearsomely difficult. Only in the present century have researchers begun to accumulate much data from fMRI imaging. New methodological problems are constantly being discovered (Vul et al., 2009; Bennett et al., 2009a, b; Button et al., 2013; Eklund et al., 2016).[a] This is not surprising given how new the science is and how difficult the problems are. Moreover, virtually no studies have been performed to address specifically normative legal questions. There are many studies of the neural correlates of legal decision-making, but they have no normative relevance. Law should not expect too much of a young science that uses new technologies to investigate some of the most intrinsically difficult problems in science and that does not directly address questions of normative interest. Caution is warranted, although many would think the argument of this chapter is too cautious.

Furthermore, neuroscience is insufficiently developed to detect specific, legally relevant mental content. For example, it does not provide a sufficiently accurate diagnostic marker for even a severe mental disorder that might be relevant to many legal doctrines (Francis, 2009; Morse and Newsome, 2013; Rego, 2016). Many studies do find differences between patients with mental disorders and controls, but the differences are too small to be used diagnostically, and publication bias may have inflated the proportion of such positive studies (Ioannides, 2011). There are limited exceptions for some genetic disorders that are diagnosed using genomic information or some well-characterized neurological disorders, such as epilepsy, that are definitively diagnosed using electroencephalography (EEG), but these are not the types of techniques that are central to the new neuroscience based primarily on imaging.

Nonetheless, certain aspects of neural structure and function that bear on legally relevant capacities, such as the capacity for rationality and control, may be temporally stable in general or in individual cases. If they are, neuroevidence may permit a reasonably valid retrospective inference, for example, about a criminal defendant's rational and control capacities and their impact

---

[a] But see Lieberman et al. (2009), p. 299.

on criminal behavior. Some legal questions, such as whether a defendant is competent and what the agent will do in the future, depend on current-state evaluation of the agent. Such an evaluation will be easier than a retrospective responsibility evaluation. Nonetheless, both types of evaluation will depend on the existence of adequate neuroscience to aid such evaluations. With few exceptions, we currently lack such science (Morse and Newsome, 2013, p. 150), but future research may provide the necessary data.

Let us consider the specific grounds for modesty about the current achievements of behavioral neuroscience. fMRI is still a rather blunt instrument to measure brain functioning. It measures the amount of oxygenated blood that is flowing to a specific region of the brain (the blood oxygen dependent level [BOLD] signal), which is a proxy for the amount of neural activation that is occurring in that region above or below baseline activation (the brain is always and everywhere physiologically active). There is good reason to believe that the BOLD signal is a good proxy, but it is only a proxy. There is a suboptimal time lag between when actual activation occurs and when fMRI measures it, and for a variety of reasons, pinpointing the exact region where activation occurred is also far from perfect (Roskies, 2013, p. 37; Hong et al., 2019). The spatial resolution problem can lead to substantial numbers of false positive findings.

These and similar technical difficulties will surely be ameliorated by technological advances, but studies to date, especially if they used lower power scanners, do suffer from these limitations. (For sophisticated, systematic and comprehensive reviews of the proper uses and limitations of fMRI, see Logothetis, 2008; Poldrack, 2018.)

There are research design difficulties. It is difficult to control for all conceivable artifacts; that is, other variables that may also produce a similar result. Consequently, there are often problems of over-inference and of invalid reverse inferences (Poldrack, 2006). The same region of interest (ROI) may be associated with opposite behaviors, which also confounds inferences.

At present, most neuroscience studies on human beings involve small numbers of subjects, which makes it difficult to achieve statistically significant results and which undermines the validity of significant findings (Button et al., 2013; Szucs and Ioannidis, 2017). This problem is termed "low power." This is especially important as research increasingly uses machine learning techniques to understand neural correlates and to make predictions. To ensure that the algorithms derived from a subject sample are generalizable, they must be cross-validated on populations different from the sample population. If sample sizes are small, the risk of error is magnified (Varoquaux, 2017). The problem of small samples will improve as the cost of scanning decreases and future studies will have more statistical power, but this is still a major problem.

The types of subjects used also present a problem. Most of the studies in cognitive, affective, and social neuroscience have been done on college and university students, who are hardly a random sample of the population generally. Many of the studies use nonhuman animals, such a rats or primates, as subjects. Although the complexity and operation of the neural structure and function of such animals may be on a continuum with those of human beings and there may be complete similarity at some level, there is reason to question the applicability of the neuroscience of behavior of nonhuman animals to humans. The human brain is capable of language and rationality, which mark an immense difference between humans and other animals. To the best of our knowledge, other animals do not act for and are not responsive to reasons in the full-blown sense that intact human beings are. Is so-called altruistic behavior in orangutans, for example, the same as altruistic behavior in humans? Although the point should not be overstated, we should be cautious about extrapolating to human action from the neuroscience of the behavior of other animals.

Most studies average the neurodata over the subjects, and the average finding may not accurately describe the brain structure or function of any individual subject in the study. This leads to a more general problem about the applicability of scientific findings from group data to an individual subject, a problem called *G2i* for "group to individual" (Faigman et al., 2014). Scientists are interested in how the world works and produce general information. Law is often concerned with individual cases, and it is difficult to know how properly to apply relevant group data. For example, a neuroscience study that reports increased activation in some brain ROI (region of interest) bases its conclusion on averaging the activation across all the subjects, but no subject's brain may have activated precisely in the area identified. If such group data are permitted, as they now are for functions such as predictions, the question is how to use probabilistic data to answer what is often a binary question, such as whether to parole a prisoner because he is deemed no longer a danger to society. This is a topic under intensive investigation at present, and I assume progress will be made.

A serious question is whether findings based on subjects' behavior and brain activity in a scanner would apply to real-world situations. This is known as the problem of "ecological validity." Does a subject's performance in a laboratory while being scanned on an executive function task that inter alia allegedly measures the

ability to control impulses really predict that person's ability to resist criminal offending, for example?

Behavioral neuroscientists have great flexibility in data collection, analysis and reporting—a phenomenon termed "researcher degrees of freedom" (Simmons et al., 2011). When this is coupled with low power (the limited potential of the study to have statistically significant findings) and the multidimensionality of fMRI data, the probability of false-positive results is increased markedly (Poldrack et al., 2017).

Replications are few, which is especially important for any discipline, such as law, that has public policy implications (Chin, 2014). Policy and adjudication should not be influenced by findings that are insufficiently established, and replications of findings are crucial to our confidence in a result, especially given the problem of publication bias (Ioannides, 2011) and reproducibility skepticism (Chin, 2014; Open Science Collaboration, 2015). Indeed, replications are so few in this young science and the power of too many studies is so low that one should be wary of the ultimate validity of many results. A recent analysis suggests that more than 50% of cognitive neuroscience studies may be invalid and not reproducible (Szucs and Ioannidis, 2017).

Drawing extended inferences from findings is especially unwarranted at present. If there are numerous studies of various types that seem valid, all converge on a similar finding, and there is theoretical reason to believe they should be consistent, then lack of replication of any one of them may not present such a large problem. For example, there are relatively few neuroscientific studies of adolescent behavior (although it is one of the areas well-studied), but they tend to be consistent with both the developmental psychology of adolescence and the neuroanatomical evidence indicating average differences between adult and adolescent brains (Galvan, 2017). But such examples are at present few, especially in legally and morally relevant neuroscience.

What is known about behavioral neuroscience is quite coarse-grained and correlational rather than fine-grained and causal (Miller, 2010). An association between a condition or a task in the scanner and brain structure or function is being investigated. These studies do not demonstrate that the brain ROI or activity is a sensitive diagnostic marker for the condition or either a necessary, sufficient, or predisposing causal condition for the behavioral task that is being performed in the scanner. Any language that suggests otherwise—such as claiming that some brain region is the "neural substrate" for the behavior—is simply not justifiable based on the methodology of most studies. Such inferences are only justified if everything else in the brain remained constant, which is seldom the case (Adolphs, 2015).

Recall the description of criminal responsibility. The law is concerned with human mental states and actions, not brain states. What is the relevance of neuroscientific evidence to decision-making concerning human behavior? If the behavioral data are not clear, then the potential contribution of neuroscience is large. Unfortunately, it is in just such cases that neuroscience at present is not likely to be of much help. I term the reason for this the "clear-cut" problem (Morse, 2011). Virtually all neuroscience studies of potential interest to the law involve some behavior that has already been identified as of interest, and the point of the study is to identify that behavior's neural correlates. Neuroscientists do not go on general "fishing" expeditions.[b] There is usually some bit of behavior—such as addiction, schizophrenia, or impulsivity—that investigators would like to understand better by investigating its neural correlates. To do this properly presupposes that the researchers have already well-characterized and validated the behavior under neuroscientific investigation. Cognitive, social, and affective neuroscience is thus inevitably embedded in a matrix involving allied sciences such as cognitive science and psychology. Behavioral science is virtually always the predicate for good behavioral neuroscience (Niv, 2021). Consequently, neurodata can seldom be more valid than the behavioral data with which it is correlated. In such cases, the neural markers might be sensitive to the already clearly identified behaviors precisely because the behavior is so clear. Less clear behavior is simply not studied, or the overlap in data about less clear behavior is greater between the subjects of interest and comparison subjects. Consequently, the neural markers of clear cases will provide little guidance to resolve behaviorally ambiguous cases of relevant behavior, and they are unnecessary if the behavior is sufficiently clear.

On occasion, the neuroscience might suggest that the behavior is not well-characterized or is neurally indistinguishable from other, seemingly different behavior. In general, however, the existence of relevant behavior will already be apparent before the neuroscientific investigation is begun. For example, some people are grossly out of touch with reality. If, as a result, they do not understand right from wrong, we excuse them because they lack such knowledge. We might learn a great deal about the neural correlates of such psychological abnormalities. But we already knew without neuroscientific data that these abnormalities existed, and we had a firm view of their normative significance.

---

[b]But see Bennett et al. (2009a,b) for an amusing exception.

In the future, we may learn more about the causal link between the brain and behavior, and studies may be devised that are more directly legally relevant. My best hope is that neuroscience and ethics and law will each richly inform the other and perhaps help reach what I term a conceptual-empirical equilibrium in some areas. I suspect that we are unlikely to make substantial progress with neural assessment of mental content, but we are likely to learn more about capacities that will bear on excuse or mitigation.

THE RADICAL CHALLENGES OF NEUROSCIENCE TO LAW

Neuroscience allegedly poses two radical challenges to current law: determinism and epiphenomenalism about mind, the no agency thesis. These are purely hypothetical, theoretical challenges at present and have virtually no practical purchase, so this chapter will deal with them briefly.

The challenge from determinism is the familiar claim that if determinism or something quite like it is true (e.g., physicalism plus causal closure), then no one can be responsible, a position termed "hard incompatibilism" (e.g., Pereboom and Caruso, 2017). Many incorrectly believe that neuroscience will prove that determinism is true, but no science can do this. It is a metaphysical hypothesis about the ontology of the known universe and, roughly, a working background hypothesis for many practicing scientists. Moreover, there is a competing position within the philosophy of responsibility, "compatibilism," which holds that we have enough freedom to ground robust responsibility even if determinism is true. This is currently the dominant position among philosophers and there is no possible resolution of this metaphysical dispute. Both camps can recognize that humans are agents who act for causal reasons, but they disagree about whether action is sufficiently free to warrant ascriptions of responsibility.

Compatibilism is the theory most consistent with the approach of the ordinary person and the law to agency and responsibility and with a scientific worldview (Morse, 2018a, b). It is entirely consistent with the moral distinctions the law makes. For example, even if determinism or something quite like it is true, some bodily movements are actions and others are not. Some defendants are deluded and most are not. Some defendants react to a do-it-or-else threat and most do not. These are simply undeniable facts about human behavior that make a moral difference on deontological and consequentialist grounds that we have good reason to accept.

There is simply no compelling reason to upend centuries of legal doctrine, theory and institutions based on an armchair metaphysical theory, hard incompatibilism, that is not itself demonstrably true and whose implications are unclear even if it were true. Given the history of the law, the burden of persuasion should rest with the radical critics of current doctrines, practices and institutions. Critics of the possibility of responsibility have a duty to provide the practical implications of their philosophical position, especially if they hope to institute radical change. They have an obligation to propose the details of psychology, politics and law that would follow. Most have not done so in any real detail. Pereboom and Caruso (2017) have admirably tried, but the "medical model" they propose, which would abandon all notions of desert, including deserved punishment, in favor of a "quarantine" model of social control, does not depend on the new neuroscience and it has been intensely criticized (Sehon, 2016; Morse, 2018a, b).

The epiphenomenal challenge is more radical. It claims that we are just a pack of neurons or victims of neuronal circumstances and that our mental states have no causal power whatsoever. On this view, minds are just the epiphenomenal foam on the neural wave. The existence of agency is thus denied, but agency is foundational for law and legal institutions. If the epiphenomenal claim is true, law and legal institutions rest on a complete illusion that is itself doing no work because illusory beliefs are mental states. Compatibilism cannot deflect this challenge because it begs the question against epiphenomenalism by assuming that we are agents, which is precisely what the radical challenge denies.

The question is whether mental epiphenomenalism is justified conceptually and empirically. Space precludes me from providing the full argument (which can be found in other work; Mele, 2009; Moore, 2012; Schurger et al., 2012; Mele, 2014; Morse, 2015; Nachev and Hacker, 2015; Schurger and Uithol, 2015), but the present conclusion, and probably the conclusion forever, is that we have no good conceptual or empirical grounds for thinking the epiphenomenal challenge is correct.

A final objection to the radical challenge is the unjustifiable normative implications that allegedly follow from the truth of the challenge. The most common is the claim, discussed previously, that the truth of the challenge implies consequentialism, a set of ethical theories that considers right action that which maximizes good consequences, and the rejection of deontology, a set of ethical theories that claims that some actions are right in themselves without regard to consequences. But if our mental states, including our reasons, are epiphenomenal and doing no work, then reasons do not have force and no normative implications follow at all. Would anyone want to live in a world without normativity (albeit the question itself makes no sense if the radical challenge is true because desires have no motivating effect but are just epiphenomena themselves)?

We are not helpless Pinocchios being dangled and manipulated by our Gepetto brains. Agency is secure, at least for now. Neuroscience will not radically transform the law's view of the person, legal doctrine and legal institutions for the foreseeable future and probably never.

LEGAL RELEVANCE

A previous section discussed the reasons to be cautious about the findings of behavioral neuroscience, but this section will assume that the scientific data being adduced to guide the law are valid. For example, it will assume that imaging data were properly acquired and interpreted. In that case, the issue will be whether the science is genuinely legally relevant. If it is not, it can be misleading and will be used primarily rhetorically.

Those who wish to understand the relevance of behavioral neuroscience to law must first understand that law is a thoroughly folk psychological institution. The primary goal of law (and morality) is to guide behavior by giving people reasons to behave one way or another (Shapiro, 2000; Sher, 2006). Law is thus like other forms of social interaction guidance and control, such as ethics, etiquette, and social norms and mores. It is addressed to creatures, us, who can understand and be guided by reasons, creatures for whom mental states in part explain their behavior. Adherents of folk psychology may differ about how mental states are individuated and guide behavior, but all are united in the belief that mental states are crucial to a full understanding of human behavior.

The criminal law is also folk psychological because virtually all its criteria for culpability, responsibility and competence are acts and mental states (Sifferd, 2006). Consequently, the crucial question for law and neuroscience is whether neuroscientific data are relevant to a legal question involving acts and mental states. Do the neuroscientific data help answer a specific legal question?

I term this the problem of "translation" (Morse, 2011). Neuroscience is a purely mechanistic science. Neuroscience eschews folk-psychological concepts and discourse (although neuroscientific articles are rife with dualistic discourse that suggests that regions of the brain are little homunculi that do things and that there seems to be a struggle between the self and the brain as an independent agent; Mudrik and Maoz, 2014). Neurons, neural networks, and the connectome do not have reasons. They have no aspirations, no sense of past, present, and future. They do not "do" things to each other. These are all properties of persons as agents. Legal rules are addressed to agents.

Is the apparent chasm between those two types of discourse bridgeable? There will always be a problem of translation between the pure mechanisms of neuroscience and the folk psychology of law. This is a familiar question in the field of mental health law (Stone, 1984), but there is even greater dissonance in neurolaw. Psychiatry and psychology sometimes treat people as mechanisms but also treat them as agents. These disciplines are in part folk psychological, and the translation will be easier than it is for neuroscience. It is the task of those doing normative neurolaw always to explain precisely how neuroscientific findings, assuming that they are valid, are relevant to a legal issue. No hand waving is allowed.

The brain does enable the mind and action (even if we do not know how this occurs). Facts we learn about brains in general or about a specific brain could in principle provide useful information about mental states and about human capacities in general and in specific cases. Some believe that this conclusion is a category error because it fallaciously attributes to a part of the person (e.g., the brain) attributes such as rationality that are logically properties of whole persons (Bennett and Hacker, 2022). This is a plausible view, and perhaps it is correct. If it is, then the whole subject of neurolaw is empty, and there was no point writing this chapter in the first place. Let us therefore bracket this pessimistic view and determine what follows from the more optimistic position that what we learn about the brain and nervous system can be potentially helpful to resolving questions of criminal responsibility and other criminal justice issues if the findings are properly translated into the law's folk psychological framework.

The question is whether some concededly valid neuroscience is legally relevant because it makes a proposition about responsibility or competence more likely to be true. Biological variables, including abnormal biological variables, do not per se answer any legal question because the law's criteria are not biological. Any legal criterion must be established independently, and biological evidence must be translated into the criminal law's folk-psychological criteria. That is, the advocate for using the data must be able to explain precisely how the neurodata bear on whether the agent acted, formed the required mens rea, or met the criteria for an excusing or mitigating condition. In the context of competence evaluations, the advocate must explain precisely how the neuroevidence bears on whether the subject was capable of meeting the law's functional criteria. If the evidence is not directly relevant, the advocate should be able to explain the chain of inference from the indirect evidence to the law's criteria. At present, few such data exist that could be the basis of such an inferential chain of reasoning (Morse and Newsome, 2013, p. 150), but neuroscience is advancing so rapidly that such data may exist in the near or medium term.

Even if neuroscience does seem relevant to a legal issue, the concerns with prejudice, cumulation (additional evidence that adds nothing new to evidence already adduced), and other worries about the potentially negative impact of concededly relevant evidence must be considered. The preliminary common wisdom about imaging data based on relatively weak studies was that they were prejudicial compared to other, equally valid sources of evidence, such as purely verbal expert testimony or psychological testing. That is, juries were likely to give brain images undue weight. More recent, better designed studies have disclosed that this worry appears unjustified. With limited exceptions, decision makers do not give undue weight to imaging data (Schweitzer et al., 2011; Roskies et al., 2013). The issue is not yet resolved empirically, but the present default should be that the evidence is not prejudicial.

A more pressing concern is the value-added of imaging. A scan is relatively expensive and somewhat time consuming. It thus has the potential for waste and delay unless there is genuine value-added. More important, legally relevant neuroimages must be based on valid, prior behavioral science that identifies clearly the behavior to which the brain structure or function will be correlated. This raises the problem of cumulation. For example, studies of the anatomical abnormalities associated with schizophrenia must have clearly identified whether the subjects in fact met the diagnostic criteria for the disorder using behavioral criteria to make the diagnosis. Thus, we already knew behaviorally that the subjects suffered from schizophrenia. What does the scan add? For another example, the law has treated adolescents differently from adults for centuries based on undoubted average behavioral differences between adolescents and adults, and recall that the criteria for responsibility are behavioral. Now we know from brain imaging data that adolescent and young adult brains are on average less anatomically mature than adult brains. What does this anatomical information add to what we already knew beyond some potentially causal information? It is unsurprising in light of the behavioral differences that there are brain differences, but would we believe adolescents are <u>not</u> behaviorally different if the current brain imaging data did not show a difference? Instead, we would justifiably believe that the neuroscience was not yet sophisticated enough to detect the undoubted brain differences.

In individual cases where the behavior is clear, the imaging data will be cumulative and unnecessary. But, might not neuroscience be especially helpful in cases in which the behavioral evidence is unclear? The answer in principle is that of course it would be helpful, but as a practical matter it will not be because the neurodata is based on correlations with clear behavioral data, a problem I described earlier as the "clear cut" issue (Morse, 2011). Where the behavior is unclear, the neurodata will not be sufficiently sensitive to help resolve the behavioral issue even if the neurodata can distinguish the already behaviorally clear cases.

Here is an example of the current limitations of neuroscience for normative conclusions. A neuroscientist and I reviewed all the behavioral neuroscience that might possibly be relevant to criminal law adjudication and policy. With the exception of a few already well-characterized medical conditions that did not employ the new neuroscience, such as epilepsy, our review found virtually no solid neuroscience findings that were yet relevant (Morse and Newsome, 2013). Similar conclusions were reached after reviews of "brain reading" studies (e.g., "neural lie detection") (Greely, 2013) and addictions (Husak and Murphy, 2013).

The generally pessimistic conclusions I have reached concerning the present relevance of imaging to behavioral legal criteria are unsurprising. Behavioral neuroscience is in its infancy and works on one of the hardest problems in science—the relationship of the brain to mental states and action. As discussed previously, most of what we know generally is correlational and coarse rather than causal and fine-grained (Miller, 2010). Replications are few, so the database necessary to reach firm conclusions simply does not exist. Virtually no study is done to address a legal question and there are problems with generalizing from the laboratory to the real world.[c] Even among populations of undoubted legal interest that have been studied intensively by neuroscientists, such as adolescents and addicts, the people who have been studied are not a random sample of the population as a whole. Future conceptual and technological advances will certainly improve our knowledge base, but, for now, modesty is in order about how neuroscience can assist with legal decision making and policy.

At this point, there has been a torrential stream of articles making inflated claims for the relevance and usefulness of neuroscience to law. I have termed this phenomenon "neurohype." Professor Francis Shen terms it, "lobbyist neuroscience," in which advocates use neuroscience more aggressively and categorically for their purposes than the evidence warrants (Shen, 2013; there are of course many examples of exemplary caution (e.g., Wax, 2016–2017), but these are less frequent especially in politically fraught legal contexts).

---

[c]But see Vilares et al. (2017) for a "proof of concept" exception.

Let us conclude this section with an observation that will always be germane even if neuroscience makes huge leaps forward. For the law, actions speak louder than images with very few exceptions. The law's criteria are behavioral—actions and mental states. If the finding of any test or measurement of behavior is contradicted by actual behavioral evidence, then we must believe the behavioral evidence because it is more direct and probative of the law's behavioral criteria except perhaps in cases of malingering (although neuroscience cannot at present reliably and validly identify malingerers).

For example, if an agent behaves rationally in a wide variety of circumstances, the agent is rational even if his or her brain appears structurally or functionally abnormal. We confidently knew that some people were behaviorally abnormal—such as being psychotic—long before there were any psychological or neurological tests for such abnormalities. In contrast, if the agent is clearly psychotic, then a potentially legally-relevant rationality problem exists even if the agent's brain looks entirely normal.

An analogy from physical medicine may be instructive. Suppose someone complains about disabling back pain, a subjective symptom, and the question is whether the subject actually does have such severe pain. We know that many people with abnormal spines do not experience back pain, and many people who complain of back pain have normal spines. If the person is claiming a disability and the spine looks dreadful, evidence that the person regularly exercises on a trampoline without difficulty indicates that there is no disability caused by back pain. If there is reason to suspect malingering, however, and there is not clear behavioral evidence of lack of pain, then a completely normal spine might be of limited use in deciding whether the claimant is malingering.

Unless the correlation between the image and the legally relevant behavior is very strong, however, such evidence will be of limited help. If a biomarker were virtually perfectly correlated with a legal criterion and it was less expensive to collect the biological data than behavioral data, then the biological variable might be a good proxy for a legal criterion. But this would be possible only with clear, bright line legal rules and not with standards, such as whether a reasonable person would be aware of a particular circumstance, because the latter have an inevitably normative component for the decision maker to assess. Further, standards can evolve, and trying to use an external marker to adjudicate them would conservatively inhibit normatively driven legal evolution. Moreover, such markers are beyond present neuroscientific expertise.

I believe that many of the claims for the relevance of neuroscience are best characterized as more "rhetorically relevant" than genuinely relevant. For example, defense advocates in capital punishment proceedings, in which the threshold for admissibility of mitigating evidence is considerably lower than at trial, hope that the fetching images produced by "real" neuroscience will be more persuasive to decision makers than evidence provided by apparently more suspect social and behavioral science, even if the advocate cannot say precisely how the neuroscience bears on a genuinely mitigating condition. Having a brain lesion or injury is not a mitigating condition per se. The actual relevance of such brain abnormality evidence therefore requires an account of why the brain evidence makes it more likely than not that a genuine mitigating condition, such as lack of rational capacity, obtains.

An instructive anecdote illustrates the point. At a conference, I was presenting to a group of federal judges the case study of Spyder Cystkopf/Herbert Weinstein, a 62-year-old retired business executive who had strangled his wife to death during an argument and then threw her out the twelfth story window of their apartment building (People v. Weinstein, 1992; Morse, 1995; Davis, 2017). It was later discovered that on the underside of the middle lining surrounding his brain, the arachnoid layer, he had a large, benign cyst that pressed on and displaced a large amount of his frontal cortex. The brain image showing the displacement is spectacularly arresting. Based on this finding, the defendant was going to raise the insanity defense, claiming that he could not conform his conduct to the requirements of the law. The behavioral history and evidence were entirely inconsistent with the validity of this claim, however, and after presentation of both the prosecution and defense arguments, 100% of the judges voted to convict. I then asked the judges if they would consider the cyst a mitigating factor at sentencing. About a third of them indicated that they would consider it, so I asked them why. The modal response was that the defendant had a proverbial "hole in his head." I asked why, if it did not affect his behavior, it should be considered a mitigating factor. None of the judges who indicated a willingness to consider it had any adequate explanation except to repeat the (true) observation that Cystkopf/Weinstein had a gross anatomical abnormality. With respect, having such an abnormality is not per se an excusing or mitigating condition unless it produces a genuine mitigating condition such as diminished rationality or diminished control capacity. But there was not a shred of evidence that the defendant had such problems. The judges simply believed that such an abnormality simply "must" have mitigating implications, but the relevance was rhetorical rather than real.

NEUROEVIDENCE IN THE CRIMINAL LAW COURTROOM

Quite recently, we finally have preliminary data about how neuroscientific information is being used in criminal cases. I do not know of any such studies in other than criminal proceedings. Six interesting empirical studies from the United States (Farahany, 2015; Gaudet and Marchant, 2016), Australia (Alimardani and Chin, 2019), England and Wales (Catley and Claydon, 2015), Canada (Chandler, 2015), and the Netherlands (De Kogel and Westgeest, 2015) have attempted to discover the extent to which and in what way neuroscientific evidence is used in criminal cases. The general interest and claims about the potential legal implications of noninvasive brain imaging by fMRI motivates this work. These studies begin to examine the reality of neuroscientific influence in criminal cases. With the exception of the Netherlands, all focus on appellate cases reported in various data bases for somewhat different periods in the range of years from 2000 to 2012, and all are admirably cautious about the extensive methodological limitations of the study sample. None purports to be an accurate representation of the use of neuroscientific evidence throughout the criminal justice system and other methodological quibbles may be raised, such as the failure to use independent inter-rater reliability (the use of independent raters to check for the accuracy of the measurement of a variable) for characterizing the cases. All use a very expansive definition of neuroscience that includes techniques and data that long antedate the new neuroscience. At most, the data are suggestive. Nonetheless, the studies are interesting and innovative.

The late, great baseball scientist Yogi Berra was apocryphally quoted as saying "It's déjà vu all over again." The data indicate that the courts make the classic mistakes about the relevance of neuroscience and behavioral genetics to criminal cases that have bedeviled the reception of behavioral science in general and of psychiatry and psychology in particular. These data did not include civil cases involving dementia, brain injury and other medical/neurological conditions in which imaging data might seem more relevant. Moreover, in many of the criminal cases, the courts did not rely solely on neuroscience.

The overarching classic mistake is misunderstanding or uncritically accepting the validity of apparently relevant science and misunderstanding the relevance of the science to the specific criminal law criteria at issue, which are primarily acts and mental states. There are no brain or nervous system criteria in criminal law for any doctrine. In particular, courts too often do not understand the following issues. Metaphysical free will is not a criterion for any criminal law doctrine, and it is not even foundational for criminal responsibility in general. Causation in general and brain causation in particular, even causation by abnormal variables, are not per se mitigating or excusing conditions, and causation per se is not the equivalent of compulsion, which is an excusing condition. And, finally, people with the same diagnosis or condition are behaviorally heterogeneous, and, ultimately, it is the behavior that is legally relevant, not the diagnosis. In one form or another, most of these cases exhibit these mistakes and confusions. It is no surprise that one of the authors, Professor Nita Farahany, characterizes the cases as follows: "That use [of neurobiological research in criminal law] continues to be haphazard, ad hoc, and often ill conceived" (Farahany, 2015, pp. 488–489).

Sentencing decisions were the most common context for the introduction of neuroscience evidence, but it was also used to resolve questions about many criminal responsibility doctrines and, surprisingly, competence, which as we have seen, is a functional behavioral determination. Perhaps the most striking finding is how infrequently the new neuroscience of functional imaging and related techniques is used. This varies across jurisdictions, but the large majority of cases involve the "old" neurology or the old neuropsychology that uses classical structural imaging or behavioral methods to assess brain functioning associated with well-characterized neurological conditions, such as epilepsy and frontal lobe injuries or lesions. Such diagnostic methods are far more common than fMRI, and, in the Dutch and Canadian samples, there is virtually no functional imaging evidence.

In sum, these studies suggest that the influence of the new neuroinvestigative techniques applied to individual cases for forensic assessment is quite modest. Even when inferences are drawn in individual cases using group data about the consequences of various neurological conditions, the studies used are often classic behavioral studies rather than neuroimaging investigations. Indeed, careful examination of the expanded case studies that the papers present indicates that, in most instances, the neuroscientific evidence was far less important than the behavioral evidence, and the former was used largely to buttress the latter. The neuroevidence was rarely dispositive, and, in the other cases, it is impossible to know from these papers' summaries of the case reports how influential the additive neuroevidence was.

The first question when considering the admissibility of scientific evidence, as always, is the degree to which the basis of the testimony has been established. We have already seen that legally relevant neuroscience is not well-established at present. For a specific example, the

apparently wide but not universal Dutch acceptance of a brain disease model of addiction that guides expert contributions to legal decision-making fails to confront the hard questions about the status of the science, even if Dutch criminal law apparently adopts the choice model of addiction. Judges are not yet in a good position to evaluate neuroscience and may be either too critical or too uncritical (see Rakoff, 2016 for an analysis by a neuroscientifically informed federal judge).

For another example, fetal alcohol syndrome (FAS) plays a large role in the Canadian cases (although not in the other samples), but the potentially legally relevant aspects of the disorder are the cognitive and rationality defects, which are behavioral signs, that sufferers demonstrate from an early age. Are the brains of FAS sufferers different from the brains of those without the disorder? Of course. This is just a necessary truth of biological materialism. If the behavior is markedly different, so will be the brain. Brain difference is not per se a mitigating or excusing condition, however. If a particular FAS sufferer is somehow sufficiently able rationally to regulate his behavior, then FAS is irrelevant to mitigation or excuse. Moreover, if a FAS sufferer exhibited lifelong cognitive defects, as many do, that sufferer is potentially excusable even if sophisticated neurotechniques cannot identify the brain pathology or brain difference.

Many of the cases in these studies fail to understand the relevance of the neuroevidence. Even if there is clear evidence of brain damage or a neurological disorder, it does not mean that the defendant did not act, lacked mens rea, was less culpable, is incompetent, or will be dangerous in the future. All the criteria depend on direct assessment of the offender's behavior. The alleged relevance of neuroevidence to competence determinations, which occurs in many of the samples, is instructive but especially bewildering. Criminal competencies are behaviorally functional, and, again, defined entirely in terms of mental states. Does the defendant understand the nature of the charges, can he rationally assist counsel, does he understand the consequence of a guilty plea, does he understand the nature of the penalty about to be imposed on him and why it is being imposed? These normative, mental criteria must all be evaluated behaviorally. Either the defendant can perform these tasks to the requisite degree or he cannot.

These are continuum capacities, however, and it may be asked whether neuroscience can help with the gray area, indeterminate cases. The answer is, no, as a result of the clear-cut problem and the heterogeneity of behavior associated with similar brain function and structure that have already been addressed.

A critical reader of the empirical studies will be repeatedly struck by how many of the expanded cases either used irrelevant or weak (or nonexistent) neuroscience—for example, to assess competence or whether a defendant suffered from a mental illness—or could have been fully resolved with more careful behavioral evaluation. Of course there can be conflict about the behavioral evidence, but because act and mental state questions must be resolved, it is the behavioral evidence that is doing the real work.

Much is at stake in criminal cases, and, of course, judges would like scientific help to resolve the vexing normative issues they must decide; but, at present, turning to the neuroscience will do nothing more in most cases than to provide a rationalization for a result the judge wishes to reach on other grounds or to avoid responsibility for having to make the hard decision directly by relying on the expert. Convergent behavioral and neurodata might help solve some of these problems that cannot be resolved with either type of evidence alone, but such convergent lines of legally relevant evidence are rare.

THE CASE FOR CAUTIOUS OPTIMISM

How can neuroscience plausibly assist the law in the near to intermediate future as the science progresses? To begin, I do not think that contemporary neuroscience raises new issues (Morse, 2004, p. 81).[d] Consider the examples of using scanning to detect mental content relevant to culpability or the possibility of using neuroscientifically-based techniques to alter the behavior of an unwilling subject. I contend that the moral, political and legal resources to address these already exist because they are simply instances of well-established theory and doctrine. The application is new, but the problems are old. There are long-standing doctrines available to address whether an investigative tool violates the Fifth Amendment protection against self-incrimination or whether the state may use techniques, such as the involuntary administration of psychotropic medication, to change an unwilling subject's behavior.

Despite having claimed that we should be exceptionally cautious about the current contributions that neuroscience can make to criminal law policy, doctrine and adjudication, I am modestly optimistic about the near and intermediate term contributions neuroscience can potentially make to our ordinary, traditional,

---

[d]But see Jones (2013) for a contrary view.

folk-psychological legal system. In other words, neuroscience may make a positive contribution even though there has been no paradigm shift in thinking about the nature of the person and the criteria for criminal responsibility. The legal regime to which neuroscience will contribute will continue to take people seriously as people—as autonomous agents who may fairly be blamed and punished based on their mental states and actions.

If a proper framework for the relevance of neuroscience to law is established and if a cautious approach to the science is adopted, I think neuroscience can potentially help refine legal mental state categories, such as mens rea and mental disorder, through a conceptual–empirical equilibrium in which legal categories guide neuroscientific investigation that in turn then help clarify the legal categories. Neuroscience might also help the fairness and efficiency of criminal law decision-making by increasing predictive accuracy. Finally, in tandem with behavioral science, neuroscience might help us more accurately understand legally relevant human capacities, such as the capacity for rationality and for self-control, which would again improve legal policy, doctrine, and adjudication. But all such optimistic outcomes will depend on precise understanding of legal relevance and valid science. Other chapters in this book may suggest contributions that would further the cautiously optimistic agenda.

More specifically, there are four types of situations in which neuroscience may be of assistance: (1) data indicating that the folk-psychological assumption underlying a legal rule is incorrect, (2) data suggesting the need for new or reformed legal doctrine, (3) evidence that helps adjudicate an individual case, and (4) data that help efficient adjudication or administration of criminal justice. Categories (3) and (4) overlap to some degree, but individual and systemic concerns can be distinguished. (In an illuminating, similar vein, Professor Owen Jones lists seven ways that neuroscience may be of assistance (Jones, 2013).)

Many criminal law doctrines are based on folk-psychological assumptions about behavior that may prove to be incorrect. If so, the doctrine should change. For example, it is commonly assumed that agents intend the natural and probable consequences of their actions. In many or most cases, it seems that they do, but neuroscience may help in the future to demonstrate that this assumption is true less frequently than we think. The question would be difficult to study neuroscientifically. Suppose, however, that a neural correlate of the formation of an intention were discovered and the study further found an unexpectedly weak connection between intentions and their probable consequences. In that case, the rebuttable presumption that people intend the natural and probable consequences of their actions that prosecutors use to prove intent should be softened or used with

more caution. As Jones points out, even if the assumption holds up, the challenge can be helpful to the law.

Second, neuroscientific data may suggest the need for new or reformed legal doctrine. For example, control tests for legal insanity have been disfavored for some decades because they are ill understood and hard to assess (American Psychiatric Association, 1983; U.S. National Institutes of Health, 2010; Gillebaart, 2018). It is at present impossible to distinguish "cannot" from "will not." Labs in many countries are working on this problem, but as yet there is no consensus. Perhaps future neuroscientific information will help to demonstrate the existence of control difficulties that are independent of cognitive incapacities. If so, then perhaps independent control tests are justified and can be rationally assessed. More generally, perhaps a larger percentage of offenders than we currently believe have such grave control difficulties that they deserve a generic mitigation claim that is not available in criminal law today. Neuroscience might help us discover that fact if a proper study to address that question could be devised. If that were true, justice would be served by adopting a generic mitigating doctrine. On the other hand, if it turns out that such difficulties are not so common, we could be more confident of the justice of current doctrine. Again, the challenge could be helpful.

Third, neuroscience might provide data to help adjudicate individual cases. Consider the insanity defense. As in *United States v Hinckley*, there is often dispute about whether a defendant claiming legal insanity suffered from a mental disorder, which disorder the defendant suffered from, and how severe the disorder was (1981, p. 1346). For example, it was unclear whether Hinckley was deluded when he tried to assassinate President Reagan and others. At present, these questions must be resolved entirely behaviorally, and there is often room for considerable disagreement about inferences drawn from the defendant's actions, including utterances. In the future, neuroscience might help resolve such questions if the clear-cut problem can be solved. Nevertheless, in the foreseeable future, I doubt that neuroscience will be able to help identify the presence or absence of specific mens reas or of specific mental content, despite some proof of concept studies (those that demonstrate that a question can be addressed or a that creation of a technique is possible, even if a study's result is not ready for practical use).

Finally, neuroscience might help us to implement current policy more efficiently. For example, the criminal justice system makes predictions about future dangerous behavior for purposes of bail, sentencing, including capital sentencing, and parole. If we have already decided that it is justified to use dangerousness predictions to make such decisions, it is hard to imagine a rational argument for doing it less accurately if we are in fact able to do it more accurately, a point raised by the President's

Commission for the Study of Bioethical Issues (Presidential Commission, 2015). Behavioral prediction techniques already exist. The question is whether neuro-scientific variables can add value by increasing the accuracy of such predictions considering the cost of gathering such data. A number of studies have been published showing the potential usefulness of neural markers for enhancing the accuracy of predictions of antisocial conduct (Aharoni et al., 2013; Pardini et al., 2014; Delfin et al., 2019; Zijlmans et al., 2021). At most, these must be considered preliminary, "proof of concept" studies. Some had sample size and content problems and none compared the value of neuroimaging to the best behavioral actuarial studies. A reanalysis of the Aharoni et al. data demonstrated that the effect size was tiny.[e] These studies are not ready to be translated into practice, but it is perfectly plausible that in the future, valid, cost–benefit justified neural markers will be identified, and thus, prediction decisions will be more accurate and just (see Poldrack et al., 2018, for a thorough review of the methodology and potential of neuroprediction).

A final example of a promising line of neuroscience research that would apply across the legal system to permit more accurate, just decision making involves the assessment of the accuracy of a subject's memory. Machine learning techniques have shown great promise in successfully retrodicting whether a subject's report of an event is accurate (Rissman et al., 2010, p. 9849; Rissman et al., 2016, p. 604). This methodology can be undermined by counter-measures subjects can employ that erase all successful retrodiction. If this problem can be surmounted, however, assessing the accuracy of witness testimony, which so often plays a crucial role in criminal and civil cases, would also be transformative for the law.

CONCLUSION

At present, neuroscience has little to contribute to more just and accurate criminal law decision-making concerning policy, doctrine and individual case adjudication. This was the conclusion reached when I tentatively first identified "Brain Overclaim Syndrome" and in "Lost in Translation," 17 years and 12 years ago respectively, and it remains true today. In the future, however, as the philosophies of mind and action and neuroscience mutually mature and inform one another, neuroscience may help us understand criminal behavior. Although no radical transformation of criminal justice is likely to occur,

neuroscience can inform the law as long as it is genuinely relevant and translated into the law's folk-psychological framework and criteria.

ACKNOWLEDGMENT

REFERENCES

Adolphs R (2015). The unsolved problems of neuroscience. Trends Cogn Sci 19: 173–175.

Aharoni E et al. (2013). Neuroprediction of future arrest. Proc Natl Acad Sci U S A 110: 6223–6228.

Alimardani A, Chin JM (2019). Neurolaw in Australia: the use of neuroscience in Australian criminal proceedings. Neuroethics 12: 255–270.

American Psychiatric Association (1983). American Psychiatric Association Statement on the insanity defense. Am J Psychiatry 140: 681–688.

Bennett MR, Hacker PMS (2022). Philosophical foundations of neuroscience, Wiley-Blackwell, Hoboken, NJ.

Bennett CM, Baird AA, Miller MB et al. (2009a). Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: an argument for proper multiple comparisons correction. J Serendipitous Unexpected Results 1: 1–5.

Bennett CM, Wolford GL, Miller MB (2009b). The principled control of false positives in neuroimaging. Soc Cogn Affect Neurosci 4: 417–422.

Button KS, Ioannidis JPA, Mokrysz C et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14: 365–376.

Caruso GD (2021). Rejecting retributivism: free will, punishment, and criminal justice, Cambridge University Press, Cambridge, U.K.

Catley P, Claydon L (2015). The use of neuroscientific evidence in the courtroom by those accused of criminal offenses in England and Wales. J Law Biosci 2: 510–549.

Chandler J (2015). The use of neuroscientific evidence in Canadian criminal proceedings. J Law Biosci 2: 550–579.

Chin JM (2014). Psychological science's replicability crisis and what it means for science in the courtroom. Psychol Public Policy Law 2: 225–238.

Cobb M (2020). The idea of the brain: the past and future of neuroscience, Basic Books, New York.

Davis K (2017). The brain defense: murder in Manhattan and the dawn of neuroscience in America's courtrooms, Penguin, New York.

De Kogel CH, Westgeest EJMC (2015). Neuroscientific and behavioral genetic information in criminal cases in the Netherlands. J Law Biosci 2: 580–605.

---

[e]For example, a re-analysis of the Aharoni et al. (2013, n. 35) by Russell Poldrack, a noted "neuromethodologist," demonstrated that the effect size was tiny (Poldrack, accessed January 11, 2018). The study used good but not use the best behavioral predictive methods for comparison.

Delfin C, Krona H, Andiné P et al. (2019). Prediction of recidivism in a long-term follow-up of forensic psychiatric patients: incremental effects of neuroimaging data. PLoS One 14: e0217127. https://doi.org/10.1371/journal.pone.0217127.

Eklund A, Nichols TE, Knutson H (2016). Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci 113: 7900–7905.

Faigman DL, Monahan J, Slobogin C (2014). Group to individual (G2i) inference in scientific expert testimony. Univ Chic Law Rev 81: 417–480.

Farahany NA (2015). Neuroscience and behavioral genetics in US criminal law: an empirical analysis. J Law Biosci 2: 485–509.

Feldman R (2009). The role of science in law, Oxford University Press, New York.

Francis A (2009). Whither DSM-V? Br J Psychiatry 195: 391–392.

Gabriel M (2017). I am not a brain: philosophy of mind for the 21st century, Wiley, Hoboken, NJ.

Galvan A (2017). The neuroscience of addiction, Cambridge University Press, Cambridge, U.K.

Gaudet LM, Marchant GE (2016). Under the radar: neuroimaging evidence in the criminal courtroom. Drake Law Rev 64: 577–661.

Gillebaart M (2018). The operational definition of self-contol. Front Psychol 9: 1231. https://doi.org/10.3389/psyg.2018.01231.

Greely HT (2013). Mind reading, neuroscience, and the law. In: SJ Morse, AL Roskies (Eds.), A primer on criminal law and neuroscience. Oxford University Press, New York.

Hong YW, Yoo Y, Han J et al. (2019). False-positive neuroimaging: undisclosed flexibility in testing spatial hypotheses allows presenting anything as a replicated finding. NeuroImage 195: 384–395.

Hunter JD, Nedelsky P (2018). Science and the good: the tragic quest for the foundations of morality, Yale University Press, New Haven and London.

Husak D, Murphy E (2013). The relevance of the neuroscience of addiction to the criminal law. In: SJ Morse, AL Roskies (Eds.), A primer on criminal law and neuroscience. Oxford University Press, New York.

Ioannides JP (2011). Excess significance bias in the literature on brain volume abnormalities. Arch Gen Psychiatry 68: 773–780.

Jones OD (2013). Seven ways neuroscience aids law. In: A Battro, S Dehaene, W Singer (Eds.), Neurosciences and the human person: new perspectives on human activities. Pontifical Academy of Sciences, Scripta Varia.

Lieberman MD, Berkman ET, Wager TD (2009). Correlations in social neuroscience aren't voodoo: a commentary on Vul et al. Perspect Psychol Sci 4: 299–307.

Logothetis NK (2008). What we can and cannot do with fMRI. Nature 453: 869–878.

Macmillan M (2000). An odd kind of fame: stories of Phineas gage, M.I.T. Press, Cambridge, MA.

McHugh PR, Slavney P (1998). The perspectives of psychiatry, 2nd edn Johns Hopkins University Press, Baltimore.

Mele AR (2009). Effective intentions: the power of conscious will, Oxford University Press, New York.

Mele AR (2014). Free: why science hasn't disproved free will, Oxford University Press, New York.

Menninger K (1968). The crime of punishment, The Viking Press, New York.

Miller GA (2010). Mistreating psychology in the decades of the brain. Perspect Psychol Sci 5: 716–743.

Miller v. Alabama, 132 S.Ct. 2455 (2012).

Moore MS (2012). Responsible choices, desert-based legal institutions, and the challenges of contemporary neuroscience. Soc Philos Policy 29: 233–279.

Moore MS (2020). Mechanical choices: the responsibility of the human machine, Oxford University Press, New York.

Morisette v. United States, 342 US 246 (1952).

Morse SJ (1995). Brain and blame. Georgetown Law J 84: 527–549.

Morse SJ (2004). New neuroscience, old problems. In: B Garland (Ed.), Neuroscience and the law: brain, mind and the scales of justice. Dana Press, New York.

Morse SJ (2007). The (non)problem of free will in forensic psychiatry and psychology. Behav Sci Law 25: 203–220.

Morse SJ (2011). Lost in translation? An essay on law and neuroscience. In: M Freeman (Ed.), Law and neuroscience. Oxford University Press, Oxford, U.K.

Morse SJ (2013). Brain overclaim redux. Law Inequality 31: 509–534.

Morse SJ (2015). Neuroprediction: new technology, old problems. Bioethica Forum 8: 128–129.

Morse SJ (2018a). The neuroscientific non-challenge to meaning morals and purpose. In: GD Caruso, F Flanagan (Eds.), Neuroxistentialism: meaning, morals, and purpose in the age of neuroscience. Oxford University Press, Oxford, U.K.

Morse SJ (2018b). Mental disorder and criminal justice. In: E Luna (Ed.), Reforming criminal justice: a report of the academy for justice bridging the gap between scholarship and reform. Academy for Justice, Phoenix, AZ.

Morse SJ, Newsome WT (2013). Criminal responsibility, criminal competence, and prediction of criminal behavior. In: SJ Morse, AL Roskies (Eds.), A primer on criminal law and neuroscience. Oxford University Press, New York.

Mudrik L, Maoz U (2014). 'Me & my brain': exposing neuroscience's closet dualism. J Cogn Neurosci 27: 211–221.

Nachev P, Hacker P (2015). The neural antecedents to voluntary action: response to commentaries. Cogn Neurosci 6: 180–186.

Niv Y (2021). The primacy of behavioral research for understanding the brain. Behav Neurosci. https://doi.org/10.1037/bne0000471. Last accessed Aug 24, 2021.

Open Science Collaboration (2015). Psychology: estimating the reproducibility of psychological science. Science 349: aac4716-1.

Pardini DA, Raine A, Erickson K et al. (2014). Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence. Biol Psychiatry 75: 73–80.

Pardo M, Patterson D (2013). Minds, brains, and law: the conceptual foundations of law and neuroscience, Oxford University Press, New York.

People v. Weinstein, 591 N.Y.S.2d 715 (1992).

Pereboom D, Caruso GD (2017). Hard-incompatibilist existentialism: neuroscience, punishment, and meaning in life. In: GD Caruso, O Flanagan (Eds.), Neuroexistentialism: meaning, morals, and purpose in the age of neuroscience. Oxford University Press, New York.

Poldrack R (2006). Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci 10: 59–63.

Poldrack RA (2018). The new mind readers: what neuroimaging can and cannot reveal about our thoughts, Princeton University Press, Princeton.

Poldrack RA, Baker CI, Durnez J et al. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat Rev Neurosci 18: 115–126.

Poldrack R, Monahan J, Imrey P et al. (2018). Predicting violent behavior; what can neuroscience add? Trends Cogn Sci 22: 111–123.

Presidential Commission for the Study of Bioethical Issues (2015). Gray matters: topics at the intersection of neuroscience, ethics and society, vol 2. Presidential Commission for the Study of Bioethical Issues, Washington, DC.

Rakoff JS (2016). Neuroscience and the law: don't rush in. In: New York Review of Books, vol. LXIII. 30–35.

Rego MD (2016). Counterpoint: clinical neuroscience is not ready for clinical use. Br J Psychiatry 208: 312–313.

Rissman J, Greely HT, Wagner AD (2010). Detecting individual memories through the neural decoding of memory states and past experience. Proc Natl Acad Sci U S A 107: 9849–9854.

Rissman J, Chow TE, Reggente N et al. (2016). Decoding fMRI signatures of real-world autobiographical memory retrieval. J Cogn Neurosci 28: 604–620.

Robinson v. California, 370 US 660 (1962).

Roskies AL (2013). Brain imaging techniques. In: SJ Morse, AL Roskies (Eds.), A primer on criminal law and neuroscience. Oxford University Press, New York.

Roskies AL, Schweitzer NJ, Saks MJ (2013). Neuroimages in court: less biasing than feared. Trends Cogn Sci 17: 99–101.

Schurger A, Uithol S (2015). Nowhere and everywhere: the causal origin of voluntary action. Rev Philos Psychol 761–778.

Schurger A, Sitt JD, Dehaene S (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. Proc Natl Acad Sci U S A 109: E2904–E2913.

Schweitzer NJ, Saks MJ, Murphy ER et al. (2011). Neuroimages as evidence in a Mens Rea defense: no impact. Psychol Public Policy Law 17: 357–393.

Sehon SR (2016). Free will and action explanation: a non-causal, compatibilist account, Oxford University Press, Oxford, United Kingdom.

Shapiro S (2000). Law, morality and the guidance of conduct. Legal Theory 6: 127–170.

Shen FX (2013). Legislating neuroscience: the case of juvenile justice. Loyola Law Rev 46: 985–1018.

Sher G (2006). Placing blame, Oxford University Press, Oxford, U.K.

Sifferd K (2006). In defense of the use of commonsense psychology in the criminal law. Law Philos 25: 571–612.

Simmons JP, Nelson LD, Simonsohn U (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychol Sci 22: 1359–1366.

Stone A (1984). Law, psychiatry, and morality, American Psychological Association, Washington, DC.

Szucs D, Ioannidis J (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biol 15: e2000797.

The Economist (2002). Open your mind. May 25, 93.

U.S. National Institutes of Health (2010). Funding Opportunity Announcement (FOA): basic research on self-regulation. Available at: https://grants.nih.gov/grants/guide/rfa-files/rfa-ag-11-010.html. Accessed 21 January 2021.

United States v Hinckley, 525 F. Supp. 1342 (DDC 1981) 1981.

Varoquaux G (2017). Cross-validation failure: small sample sizes lead to large error bars. NeuroImage 180: 68–77.

Vilares I, Wesley MJ, Ahn W-Y et al. (2017). Predicting the knowledge—recklessness distinction in the human brain. Proc Natl Acad Sci U S A 114: 3222–3227.

Vul E, Harris C, Winkielman P et al. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspect Psychol Sci 4: 274–290.

Wax AL (2016–2017). The poverty of the neuroscience of poverty: policy payoff or false promise. Jurimetrics 57: 239–287.

Wittgenstein L (1953). Philosophical investigations, The Macmillan Company, New York.

Zijlmans J, Marhe R, Bevaart F et al. (2021). The predictive value of neurobiological measures for recidivism in delinquent male young adults. J Psychiatry Neurosci 46: E271–E280.

FURTHER READING

Gilbert DT, King G, Pettigrew S et al. (2016). Comment on 'estimating the reproducibility of psychological science'. Science 351: 1037b.

Morse SJ (2019). Neuroscience and criminal law: perils and promises. In: L AlexanderK Ferzan (Eds.), The Palsgrave handbook of applied ethics and the criminal law. Palgrave Macmillan, Cham, Switzerland.

Poldrack R (2013). How well can we predict future criminal acts from fMRI data? Last accessed August 24, 2021. http://www.russpoldrack.org/search?q=aharoni.