Judging Mechanistic Neuroscience: A preliminary conceptual-analytic framework for evaluating scientific evid….

2 authors:

Jacqueline Anne Sullivan
The University of Western Ontario

31 PUBLICATIONS **139** CITATIONS

SEE PROFILE

Emily Baron
University of Toronto

2 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Philosophy of Scientific Experimentation - Neuroscience View project

Judging Mechanistic Neuroscience: A preliminary conceptual-analytic framework for

evaluating scientific evidence in the courtroom

Emily Baron[a] and Jacqueline Sullivan[b,c†]

[a]Faculty of Law, University of Toronto, 78 Queens Park, Toronto, ON M5S 2C5

Email: et.baron@mail.utoronto.ca

[b]Department of Philosophy, [c]Rotman Institute of Philosophy, Western University, 7170

Western Interdisciplinary Research Building, 1151 Richmond St., London, Ontario

CANADA N6A 5B8

Email: jsulli29@uwo.ca

†Co-authors had equivalent input and are listed in alphabetical order.

The use of neuroscientific evidence in criminal trials has been steadily increasing. Despite progress made in recent decades in understanding the mechanisms of psychological and behavioral functioning, neuroscience is still in an early stage of development and its potential for influencing legal decision-making is highly contentious. Scholars disagree about whether or how neuroscientific evidence might impact prescriptions of criminal culpability, particularly in instances in which evidence of an accused's history of mental illness or brain abnormality is offered to support a plea of not criminally responsible. In the context of these debates, philosophers and legal scholars have identified numerous problems with admitting neuroscientific evidence in legal contexts. To date, however, less has been said about the challenges of evaluating the evidence upon which integrative mechanistic explanations that bring together evidence from different areas of neuroscience are based. As we explain, current criteria for evaluating such evidence to determine its admissibility in legal contexts are inadequate. Appealing to literature in the philosophy of scientific experimentation and theoretical work in the social, cognitive and behavioral sciences, we lay the groundwork for reforming these criteria and identify some of the implications of modifying them.

Introduction

On January 3, 1999, Andrew Goldstein, a man with a history of schizophrenia, threw

Kendra Webdale to her death in front of a New York City subway train. Goldstein was

convicted by jury of second-degree murder and his plea of not criminally responsible

based on his diagnosed schizophrenia was rejected (*People v Goldstein*, 2004). The court

allowed expert witness testimony on Goldstein's history of mental illness, but denied the

admission of a positron emission tomography (PET) scan showing Goldstein to have a

"massive reduction in metabolism in the frontal lobe and the basal ganglia" (*People v*

Goldstein, 2004, p. 38), areas of the brain widely thought to underlie judgment and motor

control, respectively. This evidence was rejected because the court found that a

schizophrenia diagnosis did not "preclude per se that [the] defendant… comprehended

either the nature and consequences of his actions or that his actions were wrong" and that,

consequently, this evidence possessed no additional probative value (p. 38).

Consider the case of Grady Nelson, who brutally murdered his wife (*State of*

Florida v Grady Nelson, 2010). After a jury convicted him of murder, they were faced

with two sentencing options: a lifetime prison sentence or execution. Nelson's attorney

sought to establish via quantitative electroencephalography (QEEG) admitted in the

context of the expert testimony of a neuroscientist that his client suffered from a brain

abnormality at the time of crime to which his actions may be attributed. Nelson's lawyer

argued that the brain abnormality should serve as a basis for mitigating his punishment,

and the QEEG evidence has been attributed as playing an important role in the jury's

decision to spare Nelson's life (See for example Jones, Wagner, Faigman, & Raichle, 2013).

In another case (See Burns & Swerdlow, 2003; Redding, 2006), a forty year old male schoolteacher began making sexual advances towards his stepdaughter and soliciting prostitutes. Neuroimaging revealed a tumor large enough to disrupt normal functioning of the frontal lobe and hypothalamus. After the tumor was removed, the man's deviant sexual behavior ceased. As lawyer and psychologist Richard Redding (2006, p. 2) claims, "brain-damaged defendants are seen everyday in American courtrooms, and in many cases, their criminal behavior appears to be the product of extremely poor judgment and self-control".

Over the past 25 years, across a number of jurisdictions, courts have admitted neuroscientific evidence that speaks to defendant culpability (e.g., Farahany, 2016; Catley & Claydon, 2015; Chandler, 2015). Much of this evidence, as described in the examples above, has consisted of results from tests used to measure brain activity (e.g., EEG) and/or brain images obtained via the techniques of positron emission tomography (PET) and structural (MRI) and functional magnetic resonance imaging (FMRI). Such evidence, when deemed admissible, is presented in the context of expert scientific testimony and is used as a basis for establishing that abnormalities in defendants' brains should factor into assessments of criminal responsibility and/or sentencing decisions.

Although recent empirical work suggests that the admission of neuroscientific evidence in the courtroom has not yet impacted legal decision-making to a significant degree (For example Schweitzer, Saks, Murphy, Roskies, Sinnott-Armstrong, & Gaudet, 2011; Morse, 2017), it is important to note that much of the evidence that has been

admitted to date has been correlational rather than causal-mechanical. Until fairly

recently, details about causal pathways linking brain abnormalities to observable

behavior have been missing. Yet, progress in identifying the evolutionary (e.g., Durrant

& Ward, 2015), biological and neural mechanisms of criminal behavior and/or

psychopathological conditions that have been associated with it may be on the horizon.

For example, a recent initiative at the United States National Institute of Mental Health

(NIMH), the *Research Domain Criteria* (RDoC) *Project* (Cuthbert & Kozack, 2013), is

aimed at developing "integrative psychobiological explanations" of behavioral functions

that are disrupted in persons with mental illness. Proponents of RDoC argue that current

categories of psychopathology found in the *Diagnostic and Statistical Manual of Mental

Disorders* (DSM) are inadequate for identifying the causes of mental illness, but that a

new set of constructs aimed at individuating *valid* domains of psychological and

behavioral functioning will facilitate causal discovery. Aligned with these conceptual

changes, novel technologies like optogenetics, which allow neuroscientists to activate and

inactivate discrete populations of neurons non-invasively in non-human animals by

means of light, are positioning systems, cellular and molecular neuroscience to establish

"necessary and sufficient link[s] between neural function and behavior" (Häusser, 2014).

Technological advances in producing, detecting and assessing behavior in non-human

animal and human subjects are also facilitating the translation of cognitive and behavioral

findings from animal to human studies (e.g., Cambridge Neuropsychological Test

Automated Battery (CANTAB)). Additionally, the rise of big data and machine learning

techniques are revealing differences in the brains of healthy and diseased populations that

are proving important for predicting psychological and behavioral dysfunction in ways

that had not been previously possible (See Huys, Maia, & Frank, 2016). If evidence from these areas of neuroscience may factor into integrative mechanistic explanations of criminal behavior, it is reasonable to suppose that such evidence may find its way into legal contexts.

Yet, how should courts evaluate such evidence? Each jurisdiction uses a set of criteria to determine the admissibility of scientific evidence into the courtroom. The criteria used depend on whether a court upholds a policy of deferring to expert scientific opinion or assumes responsibility for assessing the quality of such opinion and the scientific evidence on which it is based. Following the decision in *Frye v. United States*, (1932) for example, U.S. judges distinguished valid scientific expert testimony from mere pseudoscience by deferring to what was "generally accepted in a particular field" as scientific. After the 1993 decision in the case of *Daubert v. Merrell Dow Pharmaceuticals, Inc*., judges have undertaken a more active role in assessing scientific validity, admitting scientific evidence only in instances in which "it is based on sound methods and principles" (Faigman, 2013, p. 91). We will hereafter refer to the tests set out in these cases as the *Frye* test and the *Daubert* criteria.

In this paper, we suggest that, on their own, these criteria are inadequate to assess evidence pertaining to multi-level mechanistic explanations of behavioral and psychological functions. We contend that a more detailed analysis of the research studies on which such evidence is based is needed to ensure its usefulness in legal contexts. Valuable tools for engaging in such analysis are on offer in theoretical work in the cognitive, social and behavioral sciences and the philosophy of scientific experimentation. We aim to show that such tools, when synthesized into a working

conceptual-analytic framework, may serve as an important guide to assist judges in

making adequately informed admissibility decisions about neuroscientific evidence and

aide jurors in understanding that evidence.

We begin by describing the role that neuroscientific evidence in the form of

expert testimony has and continues to play in criminal trials. We go on to describe current

criteria of admissibility and what we perceive to be their limitations. Appealing to

theoretical work in the cognitive, social and behavioral sciences and the philosophy of

scientific experimentation, we identify and define a set of important concepts for

evaluating neuroscientific evidence in order to lay the groundwork for reforming the way

it is evaluated by judges.

Mens Rea, Folk Psychology and Mechanistic Explanations

Neuroscientific evidence, by way of expert testimony, has been admitted in both civil and

criminal trials, but we restrict our focus in this paper to criminal trials. Although we focus

primarily on the Canadian justice system, the claims we make are applicable to other

jurisdictions. To prove a criminal offense in Canada, the Crown must prove beyond a

reasonable doubt that an accused person engaged in the criminally culpable act and held a

particular mental state at the time of the act. Criminal convictions require two elements:

actus reas or "guilty act" and *mens rea* or "guilty mind". While *actus reas* refers to the

voluntary criminal action or inaction—that is, the "thing done" for which the accused is

charged—*mens rea* concerns the mental state of the person found to have engaged in the

criminally culpable act (Roach, Berger, Cunliffe, & Stribopoulos, 2015). *Mens rea* refers

specifically to intent or knowledge of wrongdoing (Gooch & Williams, 2015).

Prescription of responsibility is thus distinct from the physical act or inaction. In other words, even if a court finds the accused person to have committed the unlawful deed in question, this finding alone is insufficient for a finding of criminal culpability. In Goldstein's case, for example, there was no question about the accused's actions causing Webdale's death. At issue was whether Goldstein appreciated that pushing Webdale to her death was wrong or whether he lacked adequate appreciation of wrongdoing due to his mental illness.

Scientific expert testimony is frequently used in criminal trials to establish that the accused does not meet the criterion of *mens rea* on the basis of mental disorder or defect. If a plea of not criminally responsible due to mental disorder (hereafter referred to as NCRMD[‡]) is to be successful, *mens rea* must be absent. Looking at 16(1) of the *Canadian Criminal Code* (1985), we see that people are not responsible for "…an act committed or an omission made while suffering from a mental disorder that rendered the person incapable of appreciating the nature and quality of the act or omission or of knowing that it was wrong". The NCRMD defence thus hinges upon establishing the subjective mental state of those accused of crimes.

Some legal scholars have argued that the NCMRD defense is too restrictive because the only defendants who can benefit from it are those who are incapable of understanding the wrongfulness of their actions. Yet, there are other persons who fully comprehend that their actions constitute a criminal offence but are unable to prevent those actions due to a pathological lack of control. Although these persons seem "no more deserving of punishment than those who lack such knowledge", in Canada,

---

[‡] This is in line with Canadian terminology.

legislatures and courts do not "recognize 'irresistable impulse' as a discrete branch of mental disorder defence" in part because of the difficulties associated with "distinguish[ing] between impulses that could and could not have been resisted" (Redding, 2006, p. 1).

Historically, in the context of court cases in which the NCMRD and irresistible defenses have been used, neuroscientific evidence in the form of expert testimony has made its way into the courtroom. Some philosophers (e.g., Churchland, 2006; Roskies, 2006) and legal scholars (e.g., Redding, 2006) have suggested that findings from neuroscience do bear on the question of the actions over which human beings may be said to have conscious control. They argue that by illuminating the neurobiological mechanisms of "self-control", neuroscientific evidence may speak to the question of whether a given individual who committed a crime was or was not in control of their actions. Failure to allow evidence of brain disease or disorder that may compromise an individual's self control to be admitted in criminal trials actually denies "brain-disordered defendants" a legal right, namely, "the opportunity to prove that they lacked criminal responsibility for the charged offense" (Redding 2006, 2).

Notice that the NCRMD defense and "the irresistible impulse" defense are both considered to be "mental" rather that "brain" disorder defenses. Criminal law is rooted in the conceptual-explanatory framework of folk psychology—it assumes that beliefs, desires, feelings and intentions are causally responsible for behavior and that criminal behavior may be caused in cases of mental illness by a "disordered mind" rather than a "disordered brain". Increasingly, however, neuroscience is conceiving of mental disorders as "biological disorders involving brain circuits that implicate specific domains

of cognition, emotion, or behavior" (Insel, 2013). This shift in perspective prompted the

development of NIMH's RDoC project, which puts forward a set of putatively valid

constructs designating psychological functions that may be disrupted in persons with

mental illness. In conjunction with the proposed constructs (e.g., acute threat), scientists

involved in the project have identified a set of experimental paradigms (e.g., Trier Social

Stress Test, fear-conditioning task) that they regard as appropriate for producing,

detecting and measuring those functions and identifying those entities and activities

involved in their production—units of analysis including genes, molecules, cells, circuits

and behavior. The constructs, units of analysis and paradigms are organized into a matrix,

which essentially serves as an online database into which data obtained from research

studies directed at investigating the mechanisms productive of these functions may be

organized. Research findings inputted into the matrix ultimately are supposed to yield

integrative multi-mechanistic explanations of the functions designated by the constructs

(e.g., Cuthbert, 2016). Knowledge about these mechanisms purportedly will enable

investigators to determine when they are disrupted and what the impact of such

disruption is on cognitive and behavioral functioning.

Although it is premature to speculate whether this shift in understanding mental

illness among research scientists will have any impact on the law, it is not unreasonable

to suppose that in the context of requesting evidence to be admitted in the courtroom

judges will increasingly find that a defendant's diagnosis of mental disorder is

supplemented with evidence pertaining to a disruption in multi-level mechanisms

underlying "normal" behavioral and psychological functioning. Moreover, given that

RDoC constructs have been deemed *tentatively* valid by scientists because they are

designed to individuate psychological and behavioral functions in terms of proximal underlying causes, they may be more likely to be considered valid for forensic purposes than DSM categories.

In support of this possibility, consider the case of post-traumatic stress disorder (PTSD), which has been used as a basis for criminal defenses (See for example Berger, McNiel & Binder, 2012). Historically, the defense in such trials has been required to establish that the defendant experienced a serious trauma and that the criminal act committed is causally attributable to stress caused by that trauma. Recently, neuroscientists are uncovering differences in how persons with PTSD process threatening stimuli or "acute threat" (an RDoC construct), compared to healthy controls. Emerging evidence suggests that persons with PTSD have heightened defensive responses to *subconsciously* threatening stimuli—responses over which they cannot exert conscious control (e.g., Lanius, Rabellino, Boyd, Harricharan, Frewen & McKinnon, 2017). These defensive responses could potentially constitute or result in criminal acts.

To date, evidence suggests that subcortical mechanisms involved in *subconscious threat detection*, which are part of neural circuits that comprise an evolutionarily adaptive "innate alarm system", may be altered in persons with PTSD compared to healthy controls. Components of this system, including the brainstem, amgydala, medial pre-frontal cortex, parahippocampus and visual cortex, have been observed to be hyperreactive in persons with PTSD in response "to fear or trauma related stimuli" (Lanius et al., 2017, p. 109). These mechanisms have been widely investigated in fear-conditioning paradigms in animal models (e.g., LeDoux & Pine, 2016). Thus, it is possible that in criminal cases in which PTSD is part of the defense, neuroscientific

evidence from a combination of animal and human studies might be admitted as a basis

for establishing that these regions were "hyperreactive" compared to the same regions in

healthy individuals and that non-conscious subcortical mechanisms could have resulted in

those criminal behaviors (e.g., "fight or flight" responses) for which the defendant is on

trial. Evidence used in this way would constitute a mechanistic explanation that spans the

levels of neural circuits to behavior.

It may be argued that barring the use of neuroscientific evidence, much like that

just described, as a means of establishing a mechanistic explanation for an accused's

criminal offense in the courtroom seems unjust. At the same time, however, determining

whether neuroscientific evidence is admissible and relevant to the behavior at issue in a

criminal trial is a difficult task. Progress in understanding the biological mechanisms that

give rise to behavior has and continues to involve investigators working in many different

areas of neuroscience who (a) investigate entities and activities at different levels of

organization, (b) use different methodological approaches and tools, (c) make different

assumptions about the organisms they study and (d) use different conceptual-theoretical

frameworks for understanding them. The working mechanistic explanation for

subconscious threat detection in PTSD, for example, is a byproduct of evidence obtained

from intracranial electrophysiological recordings, fMRI, behavioral, and resting state

functional connectivity studies in humans, and behavioral and cellular and molecular

studies in rodents (e.g. LeDoux & Pine, 2016). It is thus unlikely that a single technical

advisor to a judge, or a judge him/herself will have the requisite expertise to understand

evidence coming from each of these different areas of science and how it is related. Yet,

it is likely that the kind of evidence under consideration for admissibility in legal contexts

will be culled from all of these different areas. That scientists in a single field are not

necessarily experts in other scientific fields also raises an important question as to

whether evidence admitted in the context of expert testimony can be adequately evaluated

by an expert trained in only a single area of neuroscience. This highlights the problem

inherent in admitting such evidence into criminal trials: How is such evidence to be

evaluated by non-experts? In the next section, we describe the criteria judges currently

use to determine the admissibility of such evidence and some limitations of these criteria.

Neuroscientific evidence and the limitations of current admission criteria

Criminal courts most commonly admit scientific evidence by way of expert testimony

(e.g., Faigman, 2013). Expert testimony—an exception to the rule barring hearsay

evidence—is justified on the basis that the jury may not be qualified to draw proper

inferences from the presented facts due to their 'technical nature' (Stewart et al., 2016).

Despite the function of expert testimony, the courts have voiced their worries about the

use of scientific evidence in the context of trials. First, judges have expressed concern

about the prejudicial effect scientific evidence may have on the jury. Three

considerations are relevant here: (a) the jury comprehension of expert testimony, (b) the

jury's tendency to place too much weight on the testimony of experts, and (c) the jury's

susceptibility to the belief that scientific evidence is inerrant. Second, the differences

between inquiry in the trial and inquiry in science may be problematic. For example,

although trial decisions are final,[§] as Charron JA commented, "…the state of scientific

knowledge is fluid" (*R v K(A),* 1999, para 86). Because of the law's aim for consistency

---

[§]Final that is, unless new evidence is available to overturn a conviction.

and the developing nature of scientific inquiry, upholding trial decisions on the basis of

scientific research may prove problematic when the consensus about the truthfulness of

evidence is open to change in the future.

To ensure the quality of scientific research used in the courtroom, scientific

evidence is vetted by judges based on a number of considerations. Courts initially

assessed the admissibility of scientific evidence using the *Frye* test, which relied on the

general acceptance of scientific research within its own 'relevant scientific community'

(*Frye v. United States*, 1932). Following the ruling in the U.S. Supreme Court in the case

of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993 and its establishment in *J-LJ*

(2000) however, a number of jurisdictions (e.g., U.S, Canada, England and Wales) have

assigned judges the task of assessing the validity of scientific evidence (See Faigman,

2013; Edmond et al., 2008). This change reasserts judges' role as evidence 'gatekeepers'

(Dufraimont, 2017) and calls upon them to evaluate scientific evidence using several

criteria including: (a) the testability of the theory or technique used, (b) whether it has

been subject to peer review and publication, (c) the established or possible error rate, and

(d) the acceptance of the technique or theory in the scientific community (*J-LJ*, 2000, at

para 33).

Although the *Daubert* criteria appear to impose more stringent standards of

evaluation than *Frye* because judges must ensure scientific validity (rather mere general

acceptance of the evidence within its scientific contexts), it is not clear whether they do

so in a meaningful way. Scholars have pointed out a number of problems with both the

way that *Daubert* is applied in practice and the criteria themselves. First, it is not clear

how the standards should be applied. There is a lack of consistency across judicial

opinion on whether scientific research must pass all four criteria or merely meet one or two of them (Gatowski, Dobbin, Richards, Ginsburg, Merlino, & Dahir, 2001). Second, judges' ability to properly assess scientific validity has been put into question because of their lack of scientific training (Gatowksi et al., 2001; Krauss & Sales, 2001), which points to an inability to properly understand what constitutes science. This is concerning as without proper training, judges risk mistakenly regard 'junk science' as valid scientific research. For example, in a study surveying judges, Gatowski and colleagues (2001) found that although the majority of U.S. judges surveyed (88%) considered testability/falsifiability to be an important *Daubert* criterion, only 6% conveyed an adequate understanding of this concept. Scholars (e.g., Gatowksi et al., 2001) also have pointed out that the way that science is understood by judge and jury is often misguided. For example, scientific evidence, as mentioned above, may be prejudicial to the jury because of the common view that science is infallible.

As illustrated here, there are numerous problems that arise at the intersection of science and law when scientific evidence is used in criminal trials. Much recent scholarship has been devoted to addressing these issues and explicating remedies in order to bridge the gap between scientific standards and rules of evidence in law (e.g., Faigman, 2013). Few scholars, however, have addressed complex problems that arise when evidence from different areas of science is combined into mechanistic explanations of behavior and presented in the context of expert testimony. While these problems may be less obvious, illuminating them serves to provide a stronger foundation from which to develop coherent standards of admission and use. Moreover, this type of analysis is

needed to address some of the courts' concerns about the nature of scientific inquiry and

knowledge (e.g., Charron JA's concern regarding the fluidity of science).

Probing the relationship between evidence and integrative mechanistic explanations

In cases that we described in the introduction, the courts considered whether expert

scientific testimony that Goldstein and Nelson's brain abnormalities compromised

"normal" behavioral functioning, was admissible. These cases are representative of the

kinds of reasoning steps likely to be present in a request for scientific evidence to be

admitted during a trial. First, evidence in some form (e.g., images obtained by means of

PET scan, structural or functional MRI (fMRI)) is used to establish that the defendant has

a brain abnormality (e.g., a lesion, a tumor) that differs from that of persons with

"normal" or average brains. Even if this evidence may indicate that the defendant's brain

looks anatomically or structurally different from the norm, as in the case of the

schoolteacher with a brain tumor in the frontal lobe (Burns & Swerdlow, 2003),

additional evidence must be presented to establish that the brain area in question

underlies or is causally implicated in a behavioral or psychological function

corresponding to the criminal behavior for which the defendant is on trial. Such evidence

could potentially be derived from a variety of different sources including: research on

human subjects, research using animal models, studies in which machine learning

methods (e.g., pattern classification) are used to establish correlations between patterns of

brain activity and behavior, metanalyses of available research studies on humans or non-

human animals pertaining to the mechanisms of a given psychological function.

Moreover, this evidence may be presented in a variety of different representational

formats (e.g., brain images, charts and graphs, flow-diagrams/models of mechanisms). In

addition to this evidence, a case would also need to be made that the criminal behavior in

which the defendant engaged is likely attributable to the mechanisms identified –

mechanisms of which the individual was not consciously aware and over which he could

not exert control. The question that a judge immediately confronts upon considering the

admissibility of such evidence is: Does this evidence contribute in a permissible way to

the issue at hand (e.g., criminal responsibility or lack thereof)? Yet, to answer this

question, an adequate understanding of the evidence and the methods by which it was

produced is required.

Historically, philosophers of law, legal experts and judges have expressed concern

particularly about the use of brain images in the courtroom (e.g. Roskies, 2007, 2008;

Sinnott-Armstrong et al., 2008) and described a small subset of other techniques (e.g.,

epigenetics) that may yield important information in the future about the causes of human

behavior that may also be relevant in legal contexts. As more is learned about these

mechanisms, judges could be faced with a variety of different types of evidence that they

will be required to evaluate on their own or will be forced to appoint others (e.g., court

appointed experts) to evaluate on their behalf. Without an appropriate set of conceptual

tools for engaging in this task, judges and court appointed experts might make

unwarranted decisions about the admissibility of neuroscientific evidence that could

result in trial outcomes that are not aligned with the tenets or aims of the legal system.

Part of the concern of the US Supreme Court in proposing the *Daubert* criteria

was to differentiate "good science" from "junk science". From our perspective, the

criteria of *testability*, *error rate*, *peer review* and *general acceptance of evidence* among

members of the relevant scientific community, which we described above, provide an

inadequate conceptual framework for analyzing neuroscientific evidence. First, the

criteria are not accompanied by detailed definitions of the concepts to which they refer

and concepts like "validity" and "testability" are unfortunately conflated. Another

problem is that by appealing to concepts that were current in mid-20th century philosophy

of science, but ignoring more recent work to update or refine these concepts, the criteria

are out of step with current scientific practices and analytic tools available for their

assessment.

During the latter half of the 20th century in the cognitive and behavioral sciences

and philosophy of science, scholars began to put forward new conceptual tools for

understanding experimentation and the production of experimental knowledge in science

(See for example Feest & Steinle, 2016; Hacking, 1983; Radder, 2003; Shadish, Cook, &

Campbell, 2002; Slaney, 2017; Sullivan, 2009, 2015). Recently, in response to the so-

called "replication crisis" in science, more detailed criteria for evaluating the

experimental production of knowledge in science are being put forward (e.g., Ioannidis

2012; LeBel, Berger, Campbell, & Loving, 2017; Nosek, Spies, & Motyl, 2012). The

available literature offers a rich set of conceptual resources that may be used as a basis

for developing a more rigorous set of standards for evaluating neuroscientific evidence to

determine its admissibility in legal contexts. Although a discussion of *all* of the relevant

conceptual tools on offer in the philosophical and scientific literature is beyond the scope

of this paper, we undertake some preliminary work here, restricting our focus to a

discussion of the concepts of *reliability, replicability, internal validity, external validity*

and *construct validity*.[**]

In order to understand these concepts and their application, it is relevant to say

something about the *basic* structure of the kinds of experiments that may yield evidence

in support of integrative multi-level mechanistic explanations of psychological functions

that could come to factor into criminal cases. We are interested primarily in the *basic*

structure of experiments in cognitive neuroscience, systems neuroscience and cellular

and molecular neurobiology. Cognitive neuroscience aims to identify which

psychological or behavioral functions are subserved by which areas of the brain in

humans and non-human primates. Systems neuroscience and cognitive neurobiology aim

to identify the neuronal networks, synapses, cells, molecules and/or genes that bring these

functions about in non-human animals. Integrative multi-level mechanistic explanations

of the kind desired for understanding psychological and behavioral function or

dysfunction purportedly arise as evidence from these different scientific areas is

"seamlessly integrated" (e.g., Piccinini & Craver, 2011, p. 308).

One way to think about experimentation in these areas of neuroscience is as a

process (e.g., Sullivan 2009, 2015). Following Woodward (2003), we regard this process

as consisting of two basic stages: (1) *data production* and (2) *data interpretation*. In the

data production stage, an investigator poses an empirical question about a psychological

or behavioral function of interest (e.g., a question about the role of a brain area, molecule

---

[**]It is relevant to note that each term we consider has been subject to different and in some cases conflicting interpretations in the scientific and philosophical literature. Resolving such disputes is beyond the scope of this paper; we aim primarily to initiate a more detailed dialogue in the relevant literature that might serve to improve current criteria of admissibility.

or network in that psychological function). She develops an experimental design and protocol (step-by-step instructions) and selects an experimental paradigm to produce, measure and detect an instance of that function in the laboratory. This paradigm includes a set of production procedures that specifies the stimuli to be presented to each subject, how they are to be presented/arranged (e.g., spatially, temporally), and how many times each stimulus is to be presented during pre-training, training, and post-training/testing phases of the experiment. It also includes measurement procedures that specify the response variables to be measured in pre-training and post-training/testing phases of the experiment and how to measure them using tools designed for such measurement. A set of detection procedures specifies what the comparative measurements of the response variables from the different phases of the experiment must equal in order to ascribe the function of interest to the organism and/or the locus of the function to a given brain area or neuronal population. Two types of experiments are typically run in conjunction with experimental paradigms: (1) *production* experiments in which the aim is to train subjects (human or non-human animals) in the paradigm in order to assess which brain areas are involved in which cognitive functions and (2) *intervention* experiments, in which the aim is to alter the activity of an entity in the brain (e.g., a structure, network, molecule) and determine the impact on behavioral performance in the paradigm.

Once the experiment has been designed, an investigator will implement the design and protocol, taking an individual subject or group of subjects and running them through each step of the protocol. After multiple experiments have been run and enough data points for each type of experimental condition and/or manipulation have been collected, the data points are combined and analyzed statistically. The statistically analyzed data

ideally discriminates one hypothesis from a set of competing hypotheses about the function produced in the laboratory. The process of data interpretation then begins. In the first phase of this process, the investigator infers that the discriminated hypothesis is true about the effect produced in the laboratory. In the second phase, she infers that it is true of the original phenomenon that prompted the empirical question in the first place.

The first concept relevant to assessing evidence arising out of the kinds of experiments just described is *reliability*. In selecting an experimental design, protocol and experimental paradigm, an investigator ideally aims to increase the likelihood that the data production process will yield a statistically analyzed set of data that can be used to discriminate one hypothesis from a set of competing hypotheses about the effect produced in the laboratory. As we understand the term, *a data production process* is *reliable* if and only if it serves this function (e.g., Sullivan 2009, 2015). Reliability ideally functions as a normative constraint on data production; concerns about it ought to be operative throughout the data production process when an experiment is designed as well as each time an investigator implements that design when running an experiment.

Our understanding of the reliability of data production processes bears some resemblance to how epistemologist Alvin Goldman (e.g., 1993) conceives of the reliability of belief-producing processes more generally.[††] According to Goldman, a belief-producing process is reliable just so long as it has a tendency to yield more true than false beliefs. One such belief producing process that Goldman identifies as having a

---

[††] Our understanding of "reliability" also bears some resemblance to Popper's (e.g., 1992 [1959], 1979) criterion of "falsifiability" or "testability" mentioned in the *Daubert* criteria. Popper believed that what differentiated science from pseudoscience were "severe tests". Deborah Mayo provides an interpretation of the severity criterion in claiming that a *severe* test of a hypothesis "would not yield [] a passing result" if that hypothesis were indeed false (1991, p. 529).

high-tendency to be reliable is human vision. He acknowledges, however, that our visual

capacities fare better in certain environmental conditions than in others. For example, if

we form a belief about an object in dim light from a distance, that belief is less likely to

be true because vision does not work well under those conditions. If we are being

"epistemically virtuous", we aim to ensure that conditions are favorable for our belief-

producing systems to provide us with accurate information about the world. It is

reasonable to assume that scientists have at least a basic understanding of the kinds of

conditions that will facilitate or impede the reliability of a data production process. When

they aim to be epistemically virtous, they strive to meet these conditions. Being human

and thus fallible, they sometimes fall short due to undetected equipment malfunction,

errors in reasoning, errors in data collection, confirmation bias and professional and

financial pressures (e.g., Ioannidis, 2012). When evaluating neuroscientific evidence

under consideration for admissibility in criminal trials, the relevant question is: Do we

have good grounds for concluding that the data production processes by which this

evidence was produced were *reliable*?[‡‡]

This brings us to the second concept we want to consider that relates directly to

reliability, namely, *replicability*. The access a judge (or anyone outside of the

investigator(s) who conducted the study) has to the data production process in a

published research study is limited to the information contained in the methods section.

[‡‡]The other three Daubert criteria may be regarded as providing proxies for answering these questions, tying reliability assessments to whether the study has passed peer review, what the error rates of the technique(s) used in the study are and that the scientific community accepts the technique used and endorses the findings or the theory based on them. Peer review does not guarantee reliability, error rates of a technique are only one source of errors in the process of data production, and the scientific communities' acceptance of a technique in a scientific does not guarantee its reliability for establishing correlational or causal relationships (e.g., See Sullivan, forthcoming).

Although the fact that a research study has successfully passed peer review and been

published may be regarded as evidence in favor of the reliability of the data production

process, it does not guarantee it; peer review is not an infallible process and even those

scientists engaged in it may have incomplete information about the data production

processes involved in a research study they are evaluating. Recent calls for *transparency*

in science are aimed at addressing this problem. They are importantly motivated by

recognition of the importance of replication for the production of cumulative knowledge

in science. *Direct replications*, "repeat a study using methods as similar as possible to the

original study such that there is no reason to expect a different result based on current

understanding of the phenomenon" (LeBel et al., 2017, p. 5; See also Nosek et al., 2012).

That another group of scientists can use an experimental design, paradigm, protocol and

statistical analysis techniques and obtain similar results increases confidence that the

finding is not idiosyncratic to a specific context.

Popper, whose work informed the development of the *Daubert* criteria, emphasized

the importance of methodological transparency and replication in claiming "no serious

physicist would offer for publication, as a scientific discovery, [an] 'occult effect', [. . .]

for whose reproduction he could give no instructions" as "the discovery would be only

too soon rejected as chimerical, simply because attempts to test it would lead to negative

results (Popper 1959, 46)". The problem, however, is that little work in contemporary

neuroscience is aimed at direct replications of scientific findings, even though recent

initiatives suggest that this may be changing (e.g. Munafò, Nosek, Bishop, Button,

Chambers, Percie Du Sert, Simonsohn, Wagenmakers, Ware & Ioannidis, 2017). For our

purposes it is relevant to point out that when evaluating research studies put forward as

evidence in criminal trials, two relevant questions to ask are: Has the research study been

directly replicated? Were the same standards of reliability upheld in the original study

maintained in the replication(s)?[§§] These questions make explicit the implicit concerns at

the heart of the Daubert criteria that pertain to the methodology of falsifiability that

Justice Blackmun regarded as a distinguishing hallmark of science.

Two concepts also essential for evaluating scientific evidence to determine its

admissibility in court cases are internal and external validity. It is important to keep the

concepts of reliability and validity distinct; reliability is necessary for validity but not

vice versa (See Boorsboom, 2004; Sullivan 2009, 2015). As we mentioned earlier in this

section, the process of data production ideally yields a statistically analyzed set of data

that adjudicates among competing hypotheses about a phenomenon of interest. The

investigator draws inferences on the basis of this data. One set of inferences pertains to

the effect under study in the laboratory. A second set concerns the original phenomenon

that prompted the empirical question that gave rise to the experimental process in the first

place. These inferences may be valid or invalid. We regard a correlative or causal claim

about an effect (e.g., psychological function) produced in the laboratory as *internally*

valid if and only if that claim is true about that effect. It is *externally valid*, if and only if

it is applicable to or true about "circumstances of interest" (See Guala, 2003; Sullivan,

2009) outside of the laboratory. While internal validity is met when a relationship can

properly be said to exist between the obtained data and the hypothesis tested in an

---

[§§]It should be noted that in the *Daubert* decision, Justice Blackman does indicate "submission to the scrutiny of the scientific community is a component of 'good science', in part because it increases the likelihood that substantive flaws in methodology will be detected" (*Daubert v Merrell Dow Pharmaceuticals Inc*, 1993, p. 593). Of course, such flaws will not be detected if aspects of the data production process have been hidden from view.

experiment, external validity asks this question: "to what populations, settings, treatment

variables, and measurement variables can this effect be generalized?" (Campbell and

Stanley, 1963, p. 5).

When evaluating research studies put forward as evidence in criminal trials,

another relevant question is: How does the correlational or causal relationship inferred on

the basis of data obtained in research study X correspond to the behavior at issue in the

criminal trial? Some of the research studies put forward may include intervention

experiments that involve the use of animal models (e.g., rodents). Here, one relevant

question will be: is the extrapolation of a correlational or causal relationship from the one

population and context to the population of which the defendant is purportedly a member

valid? With respect to research studies involving persons of the representative population

of which the defendant is purportedly a member, one relevant question is whether the

correlational effect observed in the research study is the same correlational effect that is

true about the defendant. As other scholars have pointed out, external validity can be

assessed in at least several different ways, by looking at differences between (a) a study's

instrumentation (e.g., the way by which the manipulation was incurred), (b) a study's

sample and the population (e.g., rodents) it was intended to represent (e.g. humans), and

(c) a study's sample and another population that it was not originally formed to represent

(Kam, Wilking, & Zechmeister, 2007). Although the sample studied may not align with

the group or target that one wishes to generalize to, an extrapolation might legitimately

be made because additional information supports the appropriateness of such an inference

(See for example Steel, 2008). In such cases it is relevant to make explicit what kind of

information (assumptions, concepts, theoretical commitments) the reasoning processes

from the one context to the other involve.[***]

Social scientists have not only emphasized the importance of direct replication to cumulative progress in science, but also the relevance of conceptual replications to generalizing results beyond the laboratory, i.e., external validity. *Conceptual replications* "repeat a study using different general methodology to test whether a finding generalizes to different manipulations, measurements, domains, and/or contexts" (LeBel et al., 2017, p.5). While they are undertaken with the aim of generalizability or external validity in mind, they raise some interesting epistemological issues. First, as some scholars have noted, even subtle changes in experimental designs, paradigms, protocols and/or statistical methods may ultimately result in two research studies tapping into different phenomena and/or mechanisms, making it difficult to determine if two labs are investigating the same or different phenomena (e.g., Sullivan, 2009). It may thus be best to conceive of replications as falling along a continuum with direct replications at one end and conceptual replications that differ radically from an original research study (i.e., with respect to many parameters) on the other (LeBel et al., 2017). The farther along on the continuum we are towards radically different conceptual replications, the greater the epistemological challenges associated with relating the results of the original study to that of the purported replication.[†††]

[***]Some scholars (Mook, 1983) argue that external validity is not necessarily fundamental and that experiments in the social sciences that lack it may still provide valuable information about human behavior. We do not mean to deny this point.

[†††]It is important not to underestimate the difficulties associated with determining when two research studies using different experimental paradigms and protocols are investigating the same phenomenon or different phenomena. Treating of these issues here

A final relevant concept for our purposes is *construct validity*. Although this concept has been subject to a variety of interpretations since its introduction into the psychological literature (See for example Slaney, 2017; Messick, 1988, 1995), we adhere to a modified version of the concept as introduced by Lee Cronbach & Paul Meehl (1955) and updated in the literature on causal inference in the social sciences (Shadish, Cook & Campbell, 2002).[‡‡‡] When a scientist investigates a cognitive capacity and/or its mechanisms, she will have likely grouped together instances of what she takes to be the same capacity under a concept or construct. Examples of constructs that designate cognitive capacities include: *acute threat*, *attention*, *cognitive control* and *social communication* (to name only a handful). Such constructs originate with a concept that investigators associate with certain observations that serves as basis for theory building and experimental task/paradigm design and construction.[§§§] We can describe these constructs as "valid" insofar as they group together phenomena that correspond to *bona fide* groupings of kinds of psychological functions in the world.

An investigator will ideally aim to design an experimental paradigm (as described above) so as to produce an instance of the kind of psychological or behavioral function she intends to detect and measure. In other words, she will aim for the experimental

---

is beyond the scope of the current paper but see (Sullivan, 2016) for more detailed discussion.

[‡‡‡] In the psychological literature, the notion of "construct" has been used to pick out psychological functions directly; on this interpretation, the function is itself the construct. Constructs have also been construed as concepts that pick out psychological functions. Similarly, the concept "validity" has been applied to "constructs" as well as to tests intended to measure psychological functions. (See Slaney for extensive discussion 2017).

[§§§]Insofar as we take constructs to be concepts, our account differs from accounts that take them to be the real things under study in the laboratory. We do, however, think that the psychological functions designated by constructs are the intended targets of empirical inquiry in those areas of neuroscience of interest to us here.

paradigm to have a high degree of "construct validity". On Lee Cronbach and Paul

Meehl's original account of construct validity, it "is involved whenever a test is to be

interpreted as a measure of some attribute or quality which is not operationally defined"

(Cronbach & Meehl, 1955, p. 282). We can describe an experimental paradigm as *valid*,

if it can be used to produce, detect and measure the psychological function it was

designed to measure.

Achieving construct validity is widely acknowledged to be an iterative trial-and-

error process that has been described as involving "construct explication" and "construct

assessment" (e.g., Shadish, Cook & Campbell, 2002). These sub-processes are ideally

components of the experimental process as previously described. First, an investigator

should begin by addressing which instances of worldly phenomena should be grouped

together under the concept designating the psychological function that is the object of

empirical inquiry. Then she ought to aim to select or design an experimental paradigm

that can be used to individuate an instance of that function in the laboratory. As she

engages in the processes of data production and interpretation, she should be concerned

with whether the experimental paradigm selected is well suited for the purpose of

individuating the function or whether it should be modified or replaced and a new study

undertaken. Additionally, she should question what the data she has obtained using the

paradigm indicate about the psychological function of interest and whether the construct

designating the function should be revised so as to exclude phenomena that do not belong

in the category or to include additional phenomena that do.

We see each of these reasoning steps at work in NIMH's RDoC initiative, as

investigators have identified a preliminary set of constructs designating psychological

functions and a set of experimental paradigms to produce, detect and measure them. Yet, the scientists involved in the initiative recognize that the constructs are "heuristics" that are likely to change in response to scientific discovery. If the constructs change, new experimental paradigms will have to be identified or developed.

When evaluating research studies put forward as evidence in criminal trials, there will be a number of questions related to the issue of *construct validity*. For example: What psychological capacity was the experimental paradigm supposed to measure? Is it successful? Notice that, in asking this question, the concept of construct validity is directly related to the concept of reliability as it applies to the data production process. If an experimental paradigm fails to individuate a psychological function, then any correlational or causal hypothesis that is discriminated from a set of competing correlational or causal/mechanistic hypotheses about that function when using that paradigm will be false. An investigator may not realize this error, though, if she is confident that she has individuated a discrete function using the experimental paradigm in question. Errors surrounding construct validity are only discovered as a field makes conceptual, theoretical and explanatory progress.

Judges have expressed concerns about the admission of scientific evidence more generally, and neuroscientific evidence in particular, in legal contexts in part because scientific knowledge is not secure and the legal system requires a secure foundation. Justice Blackmun, for example, acknowledged in the *Daubert* decision that "scientific conclusions are subject to perpetual revision" whereas "law, on the other hand, must resolve disputes finally and quickly" (p. 597). It is widely accepted, however, that scientific progress is iterative. In fact, many philosophers of science have argued that

identifying the mechanisms of psychological functions is an iterative process (e.g.,

Bechtel & Richardson, 1993; Craver, 2007). Even if a scientist strives to ensure that her

experiments are reliable, the effects she observes can be replicated, the inferences made

on the basis of a set of data are internally and externally valid and the techniques that she

uses individuate a discrete phenomenon of interest succeed at that task, does not mean

that a discovery will not be made in the future that will cast this evidence in a new and

problematic light. This should not, however, be a barrier to the admissibility of

neuroscientific evidence in legal contexts. It could be argued that law is an equally

iterative process, too.

Conclusion

Law is resistant to relying on neuroscientific evidence and it is uncertain to what extent

findings from mechanistic neuroscience will ultimately influence legal decision-making.

Whether such evidence has potential to directly impact legal doctrine or indirectly

influence jury decisions by changing societal views about criminal culpability, this

resistance should not only be based on the incongruity between mechanistic explanations

of behavior and existing legal doctrine. Rather, because there is the possibility for

neuroscientific evidence to have impact (however small or in whatever ways) it is also

important to evaluate such evidence on its own merit as Morse and other scholars suggest

(Morse & Newsome, 2015). We believe that it is important for judges or (advisors to

them) to have a discrete set of conceptual-analytic tools available for evaluating scientific

evidence so as to ensure that it is not rejected out of hand and that it is adequately

analyzed before it is admitted in criminal trials. In this paper, we aimed to lay the groundwork for reforming current criteria by providing one such set of tools.

References

Bechtel, W. (2008). *Mental Mechanisms: Philosophical perspectives on cognitive neuroscience.* New York, NY: Lawrence Erlbaum.

Bechtel, W. & Richardson, R. (1993). *Discovering complexity: Decomposition and localization and strategies in scientific research.* Princeton, NJ: Princeton University Press, 1973.

Berger, O., McNiel, D. & Binder, R. (2012). PTSD as a criminal defense: a review of case law. *Journal of the American Academy of Psychiatric Law,* 40(4), 509-521.

Boorsboom, D., Cramer, A., Kievit, R., Scholten, A., & Franić, S. (2009). The end of construct validity. In R. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 135-170). Charlotte, NC: IAP Information Age Publishing.

Boorsboom, D., Mellenbergh, G. & van Heerden, J. (2004). The concept of validity. *Psychological Review,* 111(4), 1061-71.

Burns, J. & Swerdlow, R. (2003). Right orbitofrontal tumor with pedophilia symptom and constructional apraxia sign. *Archives of Neurology,* 60(3), 437-440.

Campbell, D. & Stanley, J. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally and Company.

Catley, P., & Claydon, L. (2015). The use of neuroscientific evidence in the courtroom by

those accused of criminal offenses in England and Wales. *Journal of Law and the

Biosciences*, 2(3), 510–549.

Chandler, J. (2016). The use of neuroscientific evidence in Canadian criminal

proceedings. *Journal of Law and the Biosciences*, 2(3), 550-579.

Churchland, P. (2006). The big questions: Do we have free will? Retrieved from

NewScientist.com.

Craver, C. (2007). *Explaining the brain: mechanisms and the mosaic unity of

neuroscience.* Oxford, UK: Oxford University Press.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests.

Psychological Bulletin, 52, 281–302.

Criminal Code, RSC, 1985, c. C-46.

Cuthbert, B. (2016). The NIMH research domain criteria project: toward an integrated

neuroscience of mental disorders. In T. Lehner, B. Miller, M. State (Eds.),

Genomics, circuits, and pathways in clinical neuropsychiatry (pp. 397-409).

London, UK: Elsevier.

Cuthbert, B. and Kozak, M. (2013). Constructing constructs for psychopathology: the

NIMH research domain criteria. *Journal of Abnormal Psychology*, 122, 938-937.

Daubert v. Merrell Dow Pharmaceuticals, 509 US 579 (1993).

Dufraimont, L. (2017). *New Challenges for the Gatekeeper: The Evolving Law on Expert

Evidence in Criminal Cases. (2011-2012) 58 Criminal Law Quarterly, 531.*

Durrant, R. & Ward, T. (2015). *Evolutionary criminology: towards a comprehensive

explanation of crime*. London: Elsevier.

Faigman, D. (2013). Admissibility of neuroscientific expert testimony. In S. Morse & A.

Roskies (Eds.), *A Primer on Criminal Law and Neuroscience* (pp. 89-119). Oxford,

UK: Oxford University Press.

Farahany, N. (2015). Neuroscience and behavioural genetics in US criminal law: an

empirical analysis. *Journal of Law and the Biosciences,* 2(3), 485-509.

Feest, U. & Steinle, F. (2016). Experiment. In P. Humphreys (Ed.) *The Oxford handbook

of philosophy of science.* DOI: 10.1093/oxfordhb/9780199368815.013.16

Frye v United States, 293 F 1013 (DC Cir 1923).

Gatowski, S., Dobbin, S., Richards, J., Ginsburg, G., Merlino, M., & Dahir, V. (2001).

Asking the gatekeepers: a national survey of judges on judging expert evidence in a

post-Daubert world. *Law and Human Behavior*, 25, 433-458.

Goldman, Alvin. (1993). Epistemic folkways and scientific epistemology. *Philosophical

Issues*, 3, 271-285.

Gooch, G. and Williams, M. (2015). *A dictionary of law enforcement* (2nd ed.). Oxford

Reference. DOI:10.1093/acref/9780191758256.001.0001

Guala, F. (2003). Experimental localism and external validity. *Philosophy of Science*,

Supplement, 70, 1195–1205.

Hacking, I. (1983). Representing and intervening: introductory topics in the philosophy

of natural science. Cambridge, UK: Cambridge University Press.

Häusser, M. 2014. Optogenetics: the age of light. *Nature Methods*, 11, 1012-1014.

Huys, Q., T. Maia & Frank, M.(2016). Computational psychiatry as a bridge from

neuroscience to clinical applications. *Nature Neuroscience,* 19(3), 404-13.

Ioannidis, J. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science* 7(6): 645-654.

Insel, T. (2013). Transforming diagnosis, *National Institute of Mental Health's website*. https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2013/transforming-diagnosis.shtml

Jones, O., Wagner, A., Faigman, D. & Raichle, M. (2013). Neuroscientists in court. *Nature*, 14, 730-736.

Kam, C., Wilking, J., & Zechmeister, E. (2007). Beyond the "Narrow Data Base": Another Convenience Sample for Experimental Research. *Political Behavior, 29(4)*, 415-440.

Krauss, D. & Sales, B. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy, and Law*, 7(2), 267-310.

Lanius, R., Rabellino, D., Boyd, J., Harricharan, S., Frewen, P. & McKinnon, M. (2017). The innate alarm system in PTSD: conscious and subconscious processing of threat. *Current Opinion in Psychology*, 14,109-115.

LeBel, E., Berger, D., Campbell, L., & Loving, T. (2017). Falsifiability is not optional. *Journal of Personality and Social Psychology* 113(2): 254-261.

LeDoux, J. & Pine, D. (2016). Using neuroscience to help understand fear and anxiety: a two-system framework. *American Journal of Psychiatry*, 173(11), 1083-1093.

Mayo, D. (1991). Novel evidence and severe tests. *Philosophy of Science,* 58(4), 523-552.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences

from persons' responses and performances as scientific enquiry into score meaning.
American Psychologist, 50, 741–749.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning
and consequences of measurement. In H. Wainer & H. I. Braun (Eds.),
Test validity (pp. 33–46). Hillsdale, NJ: Lawrence Erlbaum.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 379-387.

Morse, S. (2017). Criminal law and neuroscience: hope or hype? [Blog post] Retrieved
from http://www.theneuroethicsblog.com/2017/08/criminal-law-and-neuroscience-hope-or.html

Morse, S. and W. Newsome. (2015). Criminal responsibility, criminal competence, and
prediction of criminal behavior. In S. Morse and A. Roskies (Eds.)*, A primer on
criminal law and neuroscience* (pp. 150-178). Oxford, UK: Oxford University
Press.

Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers, C., Percie du Sert, N.,
Simonsohn, U., Wagenmakers, E., Ware, J., & Ioannidis, J. (2017). A manifesto for
reproducible science. *Nature Human Behavior* 1. doi:10.1038/s41562-016-0021

Nosek, B., Spies, J., & Motyl, M. (2012). Scientific utopia: II. Restructing incentives and
practices to promote truth over publishability. *Perspectives on Psychological
Science*, 7, 615-631.

People v Goldstein, 14 AD (3d) 32 (NY App Div 2005).

Piccinini, G. & Craver, C. (2011). Integrating psychology and neuroscience: functional
analysis as mechanism sketches. *Synthese*, 183(3), 283-311.

Popper, K. (1992 [1959]). *The logic of scientific discovery*. New York NY: Routledge.

Popper, K. (1979). *Objective knowledge: an evolutionary approach.* Oxford, UK: Oxford

University Press.

Radder, H. (2003). *The Philosophy of Scientific Experimentation*. Pittsburgh, PA:

University of Pittsburgh Press.

R v. J.- L. J. 2000 SCC 51[2000] 2 SCR 600.

R v. K (A) (1999), 45 OR (3d) 641, 137 CCC (3d) 225 (CA).

Redding, R. (2006). The brain disordered defendant. 56 *American University Law*

Review, 51, 1-62.

Roach, K., Berger, B., Cunliffe, E., & Stribopoulos, J. (2015). *Criminal law and*

procedure: cases and materials (11th ed.). Toronto, ON: Edmond Montgomery

Publication Limited.

Roskies, A. (2008). Neuroimaging and inferential distance. *Neuroethics,* 1, 19-30.

Roskies, A. (2007). Are neuroimages like photographs of the brain? *Philosophy of*

Science 74(5): 860-872.

Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *TRENDS in*

Cognitive Sciences, 10(9), 419-423.

Schweitzer, N., Saks, M., Murphy, E., Roskies, A., Sinnott-Armstrong, W., & Gaudet, L.

(2011). Neuroimages as evidence in a mens rea defense: No impact. *Psychology,*

Public Policy, and Law, 17(3), 357-393.

Shadish, W., Cook, T., & Campbell, D. (2002). Experimental and quasi-experimental

designs for generalized causal inference. Boston, MA: Houghton Mifflin Company.

Sinnott-Armstrong, W., Roskies, A., Brown, T., and Murphy, E. (2008). Brain images as

legal evidence, *Episteme*, 359-373.

Slaney, K. (2017). *Validating psychological constructs: historical, philosophical and practical dimensions.* London: Palgrave Macmillan.

State of Florida v Grady Nelson No. F05-00846 (11th Fla Cir Ct Dec 4 2010).

Steel, D. (2008). *Across the boundaries: Extrapolation in biology and social science.* Oxford, UK: Oxford University Press.

Stewart, H., Berger, B. L., Conifer, E., Murphy, & R., Penney, S. (2016). *Evidence: A Canadian Casebook* (4th ed.). Toronto, Ontario: Edmond Montgomery Publishing.

Sullivan, J. (Forthcoming). Optogenetics, pluralism and progress. *Philosophy of Science.*

Sullivan, J. (2016). Construct stabilization and the unity of the mind-brain sciences. *Philosophy of Science,* 83, 662-673.

Sullivan, J. (2015). Experimentation in cognitive neuroscience and cognitive neurobiology. In J. Clausen and N. Levy (Eds.), *The handbook of neuroethics* (pp. 31-47). Dordrecht, NL: Springer.

Sullivan, J. (2009). The multiplicity of experimental protocols: A challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese*, 167, 511–539.

Woodward, J. (2003). Experimentation, causal inference, and instrumental realism. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 87-118). Pittsburgh, PA: University of Pittsburgh Press.