A Neuro-Legal Lingua Franca: Bridging Law and Neuroscience on the Issue of Self-Control

Joshua W. Buckholtz
Harvard University

Valerie Reyna
Cornell University

Christopher Slobogin
Vanderbilt University Law School

Mental Health Law & Policy Journal

A Neuro-Legal Lingua Franca: Bridging Law and Neuroscience on the Issue of Self-Control

JOSHUA W. BUCKHOLTZ[*], VALERIE REYNA[**], & CHRISTOPHER SLOBOGIN[***]

ABSTRACT

Neuroscientists are rapidly adding to our understanding of human behavior. This article argues that if the law wants the full benefits of neuro-scientific knowledge, it should attempt to develop a lingua franca—a method of communication understandable to both scientists and lawyers—based on neuro-scientific concepts. As a demonstration of such an attempt, we describe in a preliminary way how the criminal law's concept of self-control might be operationalized using constructs, domains, processes and tasks familiar to neuroscientists. In the course of doing so, we stress the limits of scientific inference (particularly as it pertains to legally relevant individual-level assessment) and the fact that, despite semantic similarities, scientific constructs often do not track with its normative precepts.

TABLE OF CONTENTS

---

* Assistant Professor of Psychology and Center for Brain Science, Harvard University; Department of Psychiatry and Center for Law, Brain and Behavior, Massachusetts General Hospital.

** Professor of Human Development, Director, Human Neuroscience Institute, and Co-Director, Magnetic Resonance Imaging Facility, Cornell University.

*** Milton Underwood Professor of Law and Affiliate Professor of Psychiatry, Vanderbilt University. We thank members of the MacArthur Neuroscience Network for their comments on earlier drafts of this article. Preparation of this Article was supported, in part, by a grant from the John D. and Catherine T. MacArthur Foundation to Vanderbilt University. Its contents reflect the views of the authors, and do not necessarily represent the official views of either the John D. and Catherine T. MacArthur Foundation or the MacArthur Foundation Research Network on Law and Neuroscience (www.lawneuro.org).

> *"To identify, before the fact, those characteristics of criminal homicides and their perpetrators which call for the death penalty, and to express those characteristics in language which can be fairly understood and applied by the sentencing authority, appear to be tasks which are beyond present human ability."*
>
> McGautha v. California[1]

> *"Our ignorance of the laws of variation is profound"*
>
> Charles Darwin[2]

I. INTRODUCTION

Over the last 30 years, cognitive neuroscience has developed tools of increasing ingenuity and sophistication in an effort to illuminate the black box of the human mind. With each new advance, faith in the power to resolve ancient questions about human behavior grows. This faith is accompanied in many quarters by the hope, and in other quarters the fear, that these tools will bring into clear, quantitative, focus the beliefs, motives, memories, and capacities of individual people.

This ambivalence toward neuroscience is particularly evident in the criminal justice setting. Dozens of articles claim that brain scans can help

1. 402 U.S. 183, 204 (1971).
2. CHARLES DARWIN, THE ORIGIN OF THE SPECIES 194 (4th ed. 2003).

determine criminal responsibility or predict future criminal behavior.[3] A number of other articles counter these claims, contending that tools such as fMRI and EEG are not now, and perhaps never will be, capable of addressing most of the complex factual issues raised by criminal law.[4] A third group of scholars, comprising a wide-ranging middle ground, is willing to contemplate the possibility that, in certain subsets of cases, understanding neurobiological processes can help resolve important issues relating to the criminal law despite the limitations of current scientific knowledge.[5]

Although the debate has been lively, most of this literature has missed a crucial point. Before we can talk about whether neuroscience has anything to offer the law, lawyers and neuroscientists need a *lingua franca*—a common language that allows communication between the two groups. To a cognitive neuroscientist, legal constructs such as "premeditation," "capacity to appreciate wrongfulness," or "irresistible impulse" are, if not meaningless, far too underspecified to usefully guide scientific inquiry. They do not map onto specific mental processes, nor are they instantiated within any particular brain circuit or combination of brain circuits. They are not valid species of human mental function, so to speak.

---

3. David Eagleman, *The Brain on Trial*, THE ATLANTIC, July/August 2011, 112–115; Editorial, *Free To Choose? Modern Neuroscience is Eroding the Idea of Free Will*, THE ECONOMIST, Dec. 19, 2006, http://www.economist.com/node/8453850; Thomas Nadelhoffer & Walter Sinnott-Armstrong, *Neurolaw and Neuroprediction: Potential Promises and Perils*, 7 PHIL. COMPASS 631 (Aug. 22, 2012) http://onlinelibrary.wiley.com/doi/10.1111/j.1747-9991.2012.00494.x/full; Thomas R. Scott, *Neuroscience May Supersede Ethics and Law*, 18 SCI. & ENGINEERING ETHICS 433 (2012).

4. Uri Maoz & Gideon Yaffe, *What Does Recent Neuroscience Tell Us About Criminal Responsibility?*, 10 J. LAW & BIOSCIENCES 1, 19 (2015) ("Important neuroscientific work on self-control has emerged in recent years, although it remains uncertain how, if at all, it bears on criminal responsibility."); Sally Satel & Scott O. Lilienfeld, BRAINWASHED: THE SEDUCTIVE APPEAL OF MINDLESS NEUROSCIENCE (2013); Stephen J. Morse, *Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty*, 62 MERCER L. REV. 837, 837–38 (2011) ("Neuroscience has many things to say but not nearly as much as people would hope, especially in relation to law. At most, in the near to intermediate term, neuroscience may make modest contributions to legal policy and case adjudication.").

5. Henry T. Greely, Commentary, *What If? The Farther Shores of Neuroethics*, 18 SCI. & ENGINEERING ETHICS 439 (2012); Owen D. Jones, Joshua W. Buckholtz, Jeffrey D. Schall & Rene Marois, *Brain Imaging for Legal Thinkers: A Guide for the Perplexed*, 2009 STAN. TECH. L. REV. 5 (2009); Owen D. Jones, Anthony D. Wagner, David L. Faigman, & Marcus E. Raichle. *Neuroscientists in Court*, 14 NATURE REVIEWS NEUROSCIENCE 730 (2013); Adam J. Kolber, *Will There Be a Neurolaw Revolution*, 89 IND. L.J. 807 (2014); Adam Lamparello, *Neuroscience, Brain Damage, and the Criminal Defendant: Who Does It Help and Where in the Criminal Proceeding is It Most Relevant?*, 39 RUTGERS L. REC. 161 (2012).

Conversely, standard cognitive neuroscience terms such as "action cancellation," "task switching" or "delayed reward discounting"—each representing a valid and distinct species of cognition that can be measured with known precision and reliability—are Greek to lawyers. When law endeavors to use neuroscience tools to assess an individual's cognitive or volitional capacity, it cannot specify which, if any, of these processes is at issue. Without a common language—a lingua franca—lawyers and neuroscientists will literally be talking past one another.

Many others have written about how law can take better advantage of scientific insights.[6] In this article, we hope to contribute to that discussion by arguing that the law should begin a dialogue aimed at crafting a working lingua franca about mental states relevant to criminal justice, and that in doing so the law should be open to relying heavily on scientific concepts. Legal policy ultimately depends on normative judgments that can only be decided by legal policymakers. However, if the law's normative preferences can be framed in a common language, scientists' ability to operationalize legal concepts and produce legally relevant findings will be enhanced. Any such findings, when communicated in this shared language, will advance the goal of an empirically grounded jurisprudence. If, in contrast, the relevant legal concepts remain underspecified to a degree that prevents appropriate operationalization by scientists -- or if there is mutual agreement that some legal concepts cannot (or should not) be rendered accessible to scientific measurement -- then experts have no business providing the courts with their opinions about such matters and the courts have no business asking for them.

The following discussion sets out some preliminary thoughts on the need for a lingua franca for scientists and lawyers, using neuroscience as the relevant scientific field and criminal law as the legal domain of interest. More specifically, we examine the role of neuroscience in addressing one of the most important and contentious criminal justice issues: self-control. After a brief elaboration of the lingua franca concept, we identify the criminal law doctrines that could be said to raise self-control issues and outline cognitive processes that are relevant to the construct of self-control. We then discuss several evidentiary considerations that a lingua franca for self-control would need to take into account, and we explore several potential legal objections to a lingua franca model. While we think these objections can be countered, we also recognize, as the two quotes that begin this article suggest, that developing a lingua franca is a very difficult undertaking. Thus, we conclude with an articulation of the ideal path

---

6. *See, e.g.*, Joshua W. Buckholtz & David L. Faigman, *Promises, Promises for Neuroscience and Law*, 24 CURR. BIO. 1659 (2014); Elizabeth Mertz, *Undervaluing Indeterminancy: Translating Social Science into Law*, 60 DEPAUL L. REV. 397 (2011).

forward, a process model that would help the law to define the concept of self-control in a manner that neuroscientists can usefully operationalize. We hope that, even if a true lingua franca is impossible, such a process will more accurately identify aspects of science that serve legal principles.

II. BEYOND TRANSLATION: THE NEED FOR A LINGUA FRANCA

The law often calls upon science to help answer legal questions. In civil cases, scientists testify about the link between toxic substances and cancer, the extent of a person's physical injuries, the likelihood that employment decisions are the result of discrimination and a host of other issues. In criminal cases, scientists offer opinions not only about self-control issues, but also proffer research concerning the accuracy of eyewitnesses, assessments of whether a person presents a high risk of reoffending, or opinions about whether an offender convicted of capital murder is intellectually disabled.

Scholars who write about the law's use of science often talk about the need to translate scientific concepts into legal parlance or vice versa. At first glance, framing the issue in terms of translation seems apt, because lawyers and scientists not only use different argots but are shaped by different cultures, with differing objectives, methods of proof, and even definitions of what is a "fact."[7] Thus, just as there is a need for both a French-English and English-French dictionary, one might conclude there is a need for both a Law-Science and a Science-Law reference work.

But the concept of a lingua franca is different. A lingua franca is understood by all the parties. As is the case with Esperanto, the first attempt at creating a universal language, the words in a lingua franca can be derived by combining and modifying existing languages or can come from whole cloth. Compared to a translational model, a lingua franca model applied to the intersection of law and science could reduce confusion, sharpen legal analysis, and improve science's usefulness to the law.

Take, for instance, evidence law's use of the concept of "reliability," which to a scientist means replicability. In *Daubert v. Merrell Dow Pharmaceuticals,* the Supreme Court held that "any and all scientific testimony or evidence" must be "reliable" to be admissible.[8] While the Court noted in a footnote that it was using the word "reliable" to mean "scientific validity,"[9] a scientist reading *Daubert* and its progeny, all of

7. GARY B. MELTON ET AL., PSYCHOLOGICAL EVALUATIONS FOR THE COURTS: A HANDBOOK FOR MENTAL HEALTH PROFESSIONALS AND LAWYERS 11–15 (3d ed., Guilford 2007).

8. Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579, 589 (1993).

9. *Id*. at 595, n.12.

which continue to refer to the "reliability" of evidence, could be excused for thinking the admissibility of expert testimony rests on its consistency rather than its accuracy. In contrast to the translation approach that *Daubert* took—in essence, saying to scientists, "When we say 'reliability' we mean what you mean by 'validity'"—a lingua franca approach would simply adopt the word "validity" for legal purposes. That move would be less confusing to experts, would more precisely focus the courts and lawyers on the right issue, and would tell experts the types of information they need to provide courts assessing the admissibility of their findings.

The same kind of points can be made with respect to other evidentiary issues. *Daubert* required not only that expert testimony be reliable (i.e. valid) but also that it "fit" the facts of the case at hand and that it be "helpful" in the sense that it will add to the fact finder's knowledge.[10] Put in more scientifically palatable terms, fit might be described in terms of "external validity," and helpfulness in terms of "incremental validity." Of course, these linguistic equations might not capture the law's true meaning; even if they do, as explored more fully below, legally trained individuals might be uncomfortable with them. But if scientists and legal scholars could agree on these types of concept mappings, a lingua franca model of communication would stipulate that evidence law should adopt those terms, with the result that scientists would better understand the type of information they need to provide courts engaged in determining the admissibility of their findings.

The bottom-line concern in developing a lingua franca, then, is whether seemingly related concepts can be captured with a precision that accords both with what the law needs and with what science is able to provide. The following analysis of the law and the science of self-control begins a discussion of whether a lingua franca is possible in that setting.

III. CRIMINAL LAW AND SELF-CONTROL

In a sense, determining when a person has sufficient "self-control" to be held responsible for his or her actions is the central task of the criminal justice system. Liability for crime is premised on the idea that an offender chooses to commit the conduct that causes harm. As leading criminal law philosopher Michael Moore puts it, "[W]e are responsible for wrongs we freely choose to do, and not responsible for wrongs we lacked the freedom (capacity and opportunity) to avoid doing."[11] Similarly, Stephen Morse has written that "minimal rationality (a cognitive capacity)

10. *Id.* at 591.

11. MICHAEL S. MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 548 (Oxford Univ. Press 1997).

and minimal self-control or lack of compulsion (a volitional capacity) are the essential preconditions for responsibility".[12] If the defendant's criminal conduct is not "freely chosen," it cannot form the basis for criminal liability.

A number of specific legal doctrines attempt to define the notion of free choice. Most obviously, insanity defense doctrine is often said to be devoted to discerning which offenders are unable to control their behavior at the time of their offense. One insanity formulation, popular until relatively recently, specifically asks whether, at the time of the offense, the defendant was "substantially unable to conform his behavior to the requirements of the law" as a result of a "mental disease or defect".[13] This language, in turn, was meant to be a more expansive version of an older insanity test, sometimes dubbed the "irresistible impulse" doctrine.

In the past few decades, most states have eliminated the "volitional impairment" component of the insanity doctrine, but have retained a second insanity formulation that inquires into whether the offender, as a result of mental disease or defect, was unable to (or lacked substantial capacity to) know or appreciate the wrongfulness of the criminal conduct.[14] Although this test appears to focus on cognitive impairment, a person who does not appreciate what he or she was doing at the time of the offense or that it was wrongful could also be said to be unable to grasp the right reasons for acting, and thus lack the ability to engage in law-abiding behavior.[15]

Insanity is only the most prominent example of a criminal law doctrine that asks questions related to self-control. The "automatism" or "involuntary act" defense provides a complete excuse for offenders who can show their crime was reflexive rather than the product of a mental command to the body. For example, the defense has been successfully raised by defendants who commit harm while experiencing a seizure or while sleeping.[16]

In homicide cases, two other self-control arguments might surface. A common argument by defendants attempting to reduce a murder charge to manslaughter is that the homicide, while intentional, was committed in the "heat of passion," either as a "reasonable" response to provocation from the victim or because of some other "extreme mental or emotional distress" that is understandable in light of the defendant's situation. Impulsivity is also

---

12. Stephen J. Morse, *Excusing the Crazy: The Insanity Defense Reconsidered*, 58 S. CAL. L. REV. 777, 782 (1985).

13. MODEL PENAL CODE § 210.6 (1962).

14. MELTON et al., *supra* note 7, at 206.

15. *Cf.* Stephen Morse, *Protecting Liberty and Autonomy: Desert/Disease Jurisprudence*, 48 SAN DIEGO L. REV. 1077, 1110 (2011) ("The capacity to grasp and be guided by good reason is the heart of normative rationality.").

16. WAYNE R. LAFAVE, CRIMINAL LAW 323 (West Publishing, 5th ed. 2010).

important in reducing a charge of first degree murder—which is usually defined in terms of "premeditation" or some sort of deliberation—to one of second degree murder—which is a homicide committed with any degree of intent in the absence of provocation.[17]

For crimes other than homicide, impaired self-control is generally not a defense, unless the type of gross lack of control associated with insanity or automatism is present. Juries and judges are permitted to infer that a person intended the natural consequences of his acts.[18] However, if an offender can produce evidence suggesting the harm associated with the crime was not intended but rather was caused "recklessly" or "negligently," he or she might be convicted of a lesser offense. Conceivably, evidence that a person tends to act in an unthinking or unreflective manner could be relevant in these situations.[19]

Along the same lines, after a person has been convicted, proof that a person is relatively impulsive might reduce the sentence. For instance, in capital cases the typical "mitigating" factors—that is, factors that might lead a jury to choose a life sentence over the death penalty —are broader versions of the insanity and provocation doctrines. Under the most popular such formulation, the decision-maker is to determine whether "the capacity of the defendant to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of the law was impaired as a result of mental disease or defect," and whether "the murder was committed while the defendant was under the influence of extreme mental or emotional disturbance".[20] In contrast to the insanity defense the first mitigator does not require "significant" impairment, and in contrast to the manslaughter scenario the second provision gives mitigating effect to *any* "extreme mental or emotional disturbance," even if it is not a "reasonable" response to provocation or some other aspect of the defendant's situation. Many noncapital sentencing schemes permit judges to reduce the presumptive sentence on similar grounds.[21]

---

17. Rebecca E. Hollander-Blumoff, *Crime, Punishment and the Psychology of Self-Control,* 61 EMORY L.J. 501, 518 (2012) ("The very definition of provocation relies on a notion of the possibility of losing control of one's actions.").

18. Sandstrom v. Montana, 442 U.S. 510, 515 (1979).

19. LAFAVE, *supra* note 16, at 284 (stating that "[i]f the defendant's conduct is in fact risky, and if the risk is obvious, so that a reasonable man would realize it, we might well infer that he did not in fact realize it," but then stating "the inference cannot be conclusive, for we know that people are not always conscious of what reasonable people would be conscious of."). The Model Penal Code generally avoids negligence as a basis for liability and usually requires, at a minimum, recklessness, which it defines as a "conscious disregard [of] a substantial and unjustifiable risk." MODEL PENAL CODE § 2.02(3).

20. MODEL PENAL CODE § 210.6(4) (1962).

21. *See, e.g.,* N.C. GEN. STAT. § 15A-1340.16(e) (2013).

Despite the numerous legal doctrines that raise self-control issues, criminal law is very reticent about affording exculpatory or mitigating effect on lack-of-control grounds. In part, this reluctance is due to the plausible stance that if, as is virtually always the case except where automatisms are involved, an offender intended to carry out the conduct associated with the crime, his or her control over behavior is sufficient to satisfy basic criminal culpability requirements.[22] It is also due to a fear that, uncabined, a lack-of-control defense would spell the end of the criminal justice system, as virtually all offenders can argue either some degree of impulsivity or that their behavior was caused by factors, such as biology or upbringing, over which they had no control.[23] A third reason is practical; as is often said, the difference between an irresistible impulse and an impulse that was not resisted is hard to discern.[24] Finally, and of most interest here, the law's reluctance is due in part to the fact that, despite valiant efforts by some,[25] it has yet to devise a serviceable notion of what self-control might mean (a goal that this article argues neuroscientists might be able to help law achieve).

Ironically, the law's resistance to broad-based self-control inquiries—so powerful when criminal defendants seek mitigation—disappears in the context of preventive detention sought by the government. In *Kansas v. Hendricks*, the Supreme Court held that if the state wants to confine an individual who has already served his sentence, it must prove not only that the person presents a substantial risk of reoffending, but also that he is "dangerous beyond [his] control."[26] In these post-sentence commitment cases, it is the prosecution, not the defense, that seeks lack-of-

22. Stephen J. Morse, *Culpability and Control*, 142 U. PA. L. REV. 1587, 1610 (1992) ("I conclude that notions of loss of control are almost always parasitic upon other justifications for excuse and that the notion of loss of control unduly threatens to mislead or confuse legislators, criminal justice system participants, and the public, even in those situations in which loss of control appears to be a plausible characterization of the defendant's behavior.").

23. Meir Dan-Cohen, *Responsibility and the Boundaries of the Self*, 105 HARV. L. REV. 959, 960 (1991) ("Concessions to determinism are as inevitable as they are perilous: they are steps down a slope that, despite enormous philosophical effort, remains a slippery one.").

24. MELTON ET AL., *supra* note 7, at 216.

25. *See* Hollander-Blumoff, *supra* note 17, at 529–545 (exploring difficulties of relying "construal theory" and the "strength model of self-control" as means of constructing a legally relevant lack-of-control construct); Paul Litton, *Is Psychological Research on Self-Control Relevant to Criminal Law?*, 11 OHIO ST. J. CRIM. L. 725, 726 (2014) (arguing that the research reported by Hollander-Blumoff "does not suggest that the law underdescribes the scope of situations in which individuals could not control themselves.").

26. Kansas v. Hendricks, 521 U.S. 346, 358 (1997).

control evidence, evidence that addresses, in essence, the same issue that the volitional prong of the insanity defense describes.[27] In contrast to the paucity of lack-of-control findings when the defense proffers such evidence, they are routinely endorsed in this setting.[28]

Assume now that a defense attorney, prosecutor, or judge requests a neuroscientist to testify about a particular person's ability to exercise self-control or his capacity to "freely choose" his or her criminal conduct. The referring party is not likely to frame the issue so broadly, but rather will probably fine-tune the query by referencing the specific legal doctrine or doctrines at issue. As the foregoing discussion indicates, those doctrines ask questions regarding the defendant's capacity to: (1) conform behavior to the requirements of the law; (2) know or appreciate whether conduct is wrong; (3) control complex bodily functioning; (4) premeditate; (5) maintain composure when provoked or in other states of extreme mental or emotional stress; and (6) contemplate the consequences of his or her actions.

Can a neuroscientist provide information that is helpful in answering these inquiries? If the relevant science is lacking, do these inquiries at least tell neuroscientists what type of research they need to conduct to provide legally relevant information? Or do they require elaboration or reframing before neuroscientists can begin to offer such information or conduct the research that might produce it? Before we answer these questions, consider how neuroscientists define self-control and what they know about it so far.

IV. NEUROSCIENCE AND SELF-CONTROL

Many of us have an intuitive sense that, though we may not be able to define self-control precisely, we know it (or really, its absence) when we see it. This sensibility seems to pervade many of the capacity standards in law. In essence, each seems to ask if the defendant would not have committed the crime "but for" some very basic "deficit" in the capacity to exert "self-control." The underlying assumption is that some individuals, by quirk of brain or biology, lack this capacity, either generally or under certain circumstances. To a cognitive neuroscientist, this conceptualization poses three fundamental problems.

First, it treats self-control as if it were a unitary capacity, when in fact there are many distinct (i.e., dissociable) cognitive and socio-emotional

---

27. *Id*. at 358–59.

28. Janine Pierson, *Construing* Crane*: Examining How State Courts Have Applied Its Lack-of-Control Standard*, 160 U. PA. L. REV. 1527, 1536–37 (2012) (describing state approaches to lack-of-control issue in commitment cases).

capacities that comprise what we understand as "self-control."[29] Our folk understanding of self-control fails to account for the fact that we can observe preserved function in one capacity and poor function in another seemingly related function.[30] For example, consider a defendant who is able to value long-term goals and make appropriate decisions to achieve them (i.e. "delay gratification"), but shows deficits in the ability to learn and update associations between actions and outcomes. This defendant would be capable of prospection and future-oriented thinking in the abstract, but may not be able to forecast the bad consequences of an action that had previously resulted in harm, due to a basic deficit in the ability to use punishment-or-reward feedback to predict the outcomes of his actions. Or a person may be able to update action-outcome contingencies well, but find inhibiting his or her motor responses very challenging. As suggested above and explained in more detail below, the pertinent legal sources do not make clear which capacity or capacities is most relevant for the purpose of mitigation.

Second, while legal standards recognize some degree of variability among people, they assume that a valid, objective distinction between "able" and "deficient" or "normal" and "abnormal" individuals exists. In truth, individuals vary continuously with respect to many (if not most) cognitive capacities.[31] While some people necessarily fall at the left tail of

---

29. A considerable amount of research supports this proposition. Joshua W. Buckholtz, *Social Norms, Self-Control, and the Value of Antisocial Behavior*, 3 CURRENT OPINION BEHAV. SCI. 122 (2015); J. David Jentsch, James R. Ashenhurst, M. Catalina Cervantes, et al., *Dissecting Impulsivity and its Relationship to Drug Addictions*, 1327 ANNALS OF N.Y. ACAD. OF SCI. 1 (2014); Leigh Sharma, Kristian E. Markon, & Lee A. Clark, *Toward a Theory of Distinct Types of "Impulsive" Behaviors: A Meta-Analysis of Self-Report and Behavioral Measures,* 140 PSYCHOL. BULL. 408 (2014); Nienke Broos, Lianne Schmaal, Joost Wiskerke, et al., *The Relationship between Impulsive Choice and Impulsive Action: A Cross-Species Translational Study*. 7 PLOS ONE e36781 (2012); Damien Brevers et al., *Impulsive Action but Not Impulsive Choice Determines Problem Gambling Severity*, 7 PLOS ONE e50647 (2012); Jeffrey W. Dalley, Barry J. Everitt, & Trevor W. Robbins, *Impulsivity, Compulsivity, and Top-Down Cognitive Control*, 69 NEURON 680 (2011); Angela L. Duckworth & Margaret L. Kern, *A Meta-Analysis of the Convergent Validity of Self-Control Measures*, 45 J. RESEARCH IN PERSONALITY 259 (2011); Shashwath A. Meda, Michael C. Stevens, Marc N. Potenza, et al., *Investigating the Behavioral and Self-Report Constructs of Impulsivity Domains Using Principal Component Analysis*, 20 BEHAV. PHARMACOLOGY 390 (2009); Jeffrey W. Dalley, Adam C. Mar, Daina Economidou, & Trevor W. Robbins, *Neurobehavioral Mechanisms of Impulsivity: Fronto-striatal Systems and Functional Neurochemistry*, 90 PHARMACOLOGY BIOCHEMISTRY & BEHAV. 250 (2008); Catherine A. Winstanley, Dawn M. Eagle, & Trevor W. Robbins, *Behavioral Models of Impulsivity in Relation to ADHD: Translation Between Clinical and Preclinical Studies*, 26 CLINICAL PSYCHOL. REV. 379 (2006).

30. Broos et al., *supra* note 29, at 5–6; Meda et al., *supra* note 29, at 392.

31. Sharma et al., *supra* note 29, at 395; Duckworth & Kern, *supra* note 29, at 260.

the distribution of a given capacity, thus indicating impairment relative to the population at large, it is baseless to speak of someone as completely *lacking* any capacity. Complicating matters further, for most capacities relevant to self-control, we lack the large-scale normative datasets that would be required to make a reasonable comparison of an individual's capacity to that of the rest of the population (as we can do, for example, with I.Q.). Because various legal doctrines, such as the provocation doctrine and negligence liability, depend on whether the defendant's self-control capacities are the same as a "reasonable person's" self-control capacities, the inability to perform valid inter-individual comparisons is particularly problematic.

Third, even assuming we can identify legally relevant capacities and identify where in the population distribution a given individual falls with respect to those capacities, that information would not answer the normative question of whether a person crosses the threshold into "deficient" or "abnormal." Neuroscience and cognitive science research can examine the relations between capacities and various behaviors such as substance abuse or criminal acts. Ultimately, however, the law must determine what degree of deviation from the population mean of a given capacity is legally relevant. Further, if this decision is made in the absence of compelling scientific data, the law must recognize the arbitrary nature of any such cut-point.

With these caveats in mind, we outline below a basic neuroscientific framework for thinking about self-control that might be useful in addressing the legal standards discussed above. Cognitive neuroscientists often consider relationships among brain, mind, and behavior in terms of "constructs," "domains," "processes," and "tasks." Here, we offer some working definitions. *Constructs* refer to concepts that cannot be directly observed, but that plausibly describe a phenomenon of interest—here self-control. *Domains* reflect distinct branches or subdivisions of a construct; the self-control domains considered in this article are impulsive action, impulsive choice, and behavioral flexibility. Within a domain, there may be several *processes*, which can be thought of as specific types of mental operations or computations. Within the domain of impulsive action, for instance, neuroscientists might conceive of several processes, including motor response cancellation and motor response suppression. Finally, such processes can be measured through performance on *tasks*—experimental paradigms that are designed to index specific cognitive processes (or domains of processes)—although, as made clear below, process-task mappings are often fuzzy.

A. *Self-Control Domains: Action, Choice, and Flexibility*

Behavioral and neurobiological data suggest that self-control is not a unitary construct. Lesion studies, pharmacological manipulations, and idiographic approaches all point to the notion that self-control can be parsed into several domains.[32] Cognitive capacities or processes within one domain appear to be distinct from those in other domains. In other words, a given individual can show deficits in some self-control processes and preserve functioning in others. Although we currently lack the extensive empirical foundation required to articulate a definitive taxonomy of self-control,[33] available data highlight three legally relevant self-control domains: action, choice, and behavioral flexibility.[34] Here, we refer to self-control deficits in these three domains as "impulsive action," "impulsive choice," and "behavioral inflexibility," respectively.

1. Impulsive Action

The self-control domain of action involves processes that enable people to use external cues to inhibit "pre-potent" (habitual or dominant) motor responses. For example, "action suppression" is thought to support the ability to prevent the generation of a motor response when an external cue indicates that it is no longer appropriate.[35] This process can be tested through a task—called a Go/NoGo task—that requires execution or suppression of a motor response (action) depending on current visual input.

---

32. Dalley et al., *Neurobehavioral*, *supra* note 29, at 252; Duckworth & Kern, *supra* note 29, at 260; Meda et al., *supra* note 29, at 396–98; Sharma et al., *supra* note 29, at 393–95.

33. Robert M. Bilder et al., *Cognitive Ontologies for Neuropsychiatric Phenomics Research*, 14 COGNITIVE NEUROPSYCHIATRY 419, 429–30 (2009); Joshua W. Buckholtz & Andreas Meyer-Lindenberg, *Psychopathology and the Human Connectome: Toward a Transdiagnostic Model of Risk for Mental Illness*, 74 NEURON 990, 998–99 (2012); Russell A. Poldrack et al., *The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience*, 5 FRONTIERS NEUROINFORMATICS 2 (2011).

34. Dalley et al., *Impulsivity*, *supra* note 29, at 680; Dalley et al., *Neurobehavioral*, *supra* note 29, at 251–52; Winstanley et al., *supra* note 29, at 385.

35. *See generally*, Adam R. Aron, *From Reactive to Proactive and Selective Control: Developing a Richer Model for Stopping Inappropriate Responses*, 69 BIOLOGICAL PSYCHIATRY 55 (2011); Adam R. Aron & Russell A. Poldrack, *The Cognitive Neuroscience of Response Inhibition: Relevance for Genetic Research in Attention-Deficit/Hyperactivity Disorder*, 57 BIOLOGICAL PSYCHIATRY 1285 (2005); S. Groman, A. James & J. Jentsch, *Poor Response Inhibition: At the Nexus Between Substance Abuse and Attention Deficit/Hyperactivity Disorder*, 33 NEUROSCIENCE & BIOBEHAVIORAL REVIEWS 690 (2009).

In Go-No Go tasks, subjects are asked to view a stream of rapidly presented visual stimuli (e.g. one item every 500 milliseconds) on a computer screen. In one such task, subjects are instructed to press a button each time they see a consonant, and to avoid pressing the button each time they see a vowel. Assuming random sampling from the English alphabet, subjects will see many more consonants than vowels (a roughly four-to-one ratio) and thus will rapidly develop a tendency to press the consonant button. As a result, avoiding pressing a button when a vowel appears requires inhibition of a pre-potent motor response ("press"), and becomes relatively effortful.

In this task, the inputs might be: 1) visual stimuli (the letters on the computer), 2) a mental representation of the rules governing action ("press button for consonants, but not for vowels"), and 3) a representation of the reinforcement value of correct task performance (e.g. monetary incentives, reputation, etc.). These inputs are combined by some sort of (not yet fully characterized) computational algorithm, which sends an output that biases the selection of a specific action from an array of possible response options ("press button," "don't press button"). Individuals who have difficulty suppressing a button-press when vowels appear could be said to have reduced self-control, at least with respect to the dimension of self-control that encompasses the process of motor response inhibition. Similarly, in what is known as an anti-saccade task, subjects are asked to look in the direction that is the opposite of a briefly flashed stimulus. This ability usually improves from childhood to adulthood, and difficulty in suppressing reflexive saccades has been associated with risk taking behavior.[36]

A second process within this domain, often called "action cancellation," is thought to be crucial to the ability to use new information from the environment to inhibit the execution of a motor response once it has been initiated.[37] In contrast to response inhibition, which involves suppressing an action or response prior to its selection and initiation, action cancellation ("stopping") requires suppressing an action after its initiation. Action cancellation is often measured using a "Stop-Signal" task in which participants are asked to perform a visual discrimination (e.g. "circle or square") via a button press. On some trials ("Stop" trials), an auditory tone follows presentation of the visual cue; the tone signals that the subject must

36. Beatriz Luna, Aarthi Padmanabhan & Charles Geier, *The Adaptive Adolescent Sensation Seeking Period: Development of Reward Processing and its Effects on Cognitive Control*, THE NEUROSCIENCE OF RISKY DECISION MAKING 46 (V. Reyna & V. Zayas eds., 2014).

37. Adam R. Aron, *The Neural Basis of Inhibition in Cognitive Control*, 13 NEUROSCIENTIST 214 (2007); Adam R. Aron & Frederick Verbruggen, *Stop The Presses: Dissociating a Selective From a Global Mechanism for Stopping*, 19 PSYCHOL. SCI. 1146 (2008); Frederick Verbruggen & Gordon D. Logan, *Response Inhibition in the Stop-Signal Paradigm*, 12 TRENDS COGNITIVE SCI. 418 (2008).

refrain from making a button press. Thus, subjects must generate a button press response when they view the shape, then cancel that response if they hear the stop cue; those who are slow in cancelling the response could be said to demonstrate less self-control.

2. Impulsive Choice

From decisions about life insurance plans and 401k investments to choosing what (or what not) to eat for dinner, humans are constantly faced with choices that require weighing the benefits of a decision against its costs.[38] Impulsive choice refers to a deficit in the ability to appropriately weigh the costs, benefits, and consequences of one's actions when making a decision. Processes within this domain may be complicated by the nature of the costs and benefits that need to be integrated in order to make a decision.[39] Delay (time), probability, and effort are examples of distinct costs. Take, for example, a choice between receiving $5 and $10. In the absence of any additional cost information, the choice is straightforward: most (if not all) people would choose the higher magnitude option. But what if the choice was between receiving $5 now or $10 in 6 months; or between a 100% probability of getting $5 and a 50% probability of getting $10; or between doing 1 sit-up for $5 and 100 sit-ups for $10? Deciding between the two options in each case requires integrating information about the benefits of each choice (i.e. their absolute monetary values) with information about their costs (delay, probability, and effort, respectively).[40]

Depending on the particular kind of cost, impulsive choice failures lead to problems with delaying gratification (i.e. greater sensitivity to immediate rewards or to the cost of delaying a reward) or to increased risk-taking (i.e. greater sensitivity to the prospect of greater rewards or lower sensitivity to the prospect of bad outcomes). The capacity to delay gratification is commonly assessed using inter-temporal choice tasks, in

---

38. David J. Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 Q. J. ECON. 443 (1997); Gregory S. Berns, David Laibson, & George Loewenstein, *Intertemporal Choice–Toward an Integrative Framework*, 11 TRENDS COGNITIVE SCI. 482 (2007).

39. Mark E. Walton et al., *Calculating the Cost of Acting in Frontal Cortex*, 1104 ANNALS N.Y. ACAD. SCI. 340 (2007); Peter H. Rudebeck et al., *Separate Neural Pathways Process Different Decision Costs*, 9 NATURE NEUROSCIENCE 1161 (2006); Alex Pine et al., *Dopamine, Time, and Impulsivity in Humans*, 30 J. NEUROSCIENCE 8888 (2010); Jan Peters & Christian Büchel, *The Neural Mechanisms of Inter-Temporal Decision-Making: Understanding Variability*, 15 TRENDS COGNITIVE SCI. 227 (2011); Stan B. Floresco et al., *Cortico-Limbic-Striatal Circuits Subserving Different Forms Of Cost-Benefit Decision Makin*, 8 COGNITIVE, AFFECTIVE & BEHAV. NEUROSCIENCE 375 (2008); Franziska Denk et al., *Differential Involvement of Serotonin and Dopamine Systems in Cost-Benefit Decisions About Delay or Effort*, 179 PSYCHOPHARMACOLOGY 587 (2005).

40. Walton et al., *supra* note 39, at 350; Floresco et al., *supra* note 39, at 386.

which subjects indicate a preference for receiving sooner-but-smaller versus larger-but-later rewards, usually varying amounts of money. Outcome variables include the proportion of "sooner" reward choices, as well as the extent to which subjects discount the value of the reward as it recedes further into the future.[41] Adaptive behavior requires an ability to inhibit risk taking or to delay gratification in order to achieve greater overall gains (taking the large sum of money later). Impulsive individuals show steeper discounting of delayed rewards; the value of a given reward to an impulsive person drops steeply as the time to receiving that reward gets further and further into the future.[42] Similarly, risk-based decision-making and ambiguous decision-making can be measured by requiring subjects to "gamble" between choice options that vary in the probability or uncertainty, respectively, of rewards and/or punishments.

3. Behavioral Inflexibility

A third domain of self-control—flexibility—includes processes that enable an individual to marshal their attentional resources to achieve a goal, particularly in the face of interference, and to adapt their decisions to changing rules and dynamic feedback.[43] The interference suppression process refers to the capacity to suppress the influence of distracting information while focusing attention to perform a task.[44] For example, in an Eriksen Flanker task, participants are asked to indicate the direction of central arrow that is "flanked" on both sides by other arrows. The flanking arrows either point in the same direction as the central arrow (congruent condition) or in the opposite direction (incongruent condition). Correct responding in the incongruent condition necessitates suppressing the interfering information provided by the flankers. Similarly, in Stroop tasks participants are shown color words written in either the same color as the

41. Peters & Büchel, *supra* note 39, at 228.

42. *Id.;* John Monterosso, Payam Piray, & Shan Luo, *Neuroeconomics and the Study of Addiction*, 72 BIOLOGICAL PSYCHIATRY 107 (2012).

43. R. Cools & M. D'Esposito, *Inverted-U-Shaped Dopamine Actions on Human Working Memory and Cognitive Control*, 69 BIOLOGICAL PSYCHIATRY 113 (2011); R. Cools, *Dopaminergic Modulation of Cognitive Function-Implications for L-DOPA Treatment in Parkinson's Disease*, 30 NEUROSCIENCE & BIOBEHAVIORAL REV. 1 (2006); R.E. Laughlin et al., *Genetic Dissection of Behavioral Flexibility: Reversal Learning in Mice*, 69 BIOLOGICAL PSYCHIATRY 1109 (2011); Catherine A. Winstanley et al., *Insight Into the Relationship Between Impulsivity and Substance Abuse from Studies Using Animal Models*, 34 ALCOHOLISM, CLINICAL & EXPERIMENTAL RES. 1306 (2010); D.G. Ghahremani et al., *Neural Components Underlying Behavioral Flexibility in Human Reversal Learning*, 20 CEREBRAL CORTEX 1843 (2010).

44. E.K. Miller & J.D. Cohen, *An Integrative Theory of Prefrontal Cortex Function*, 24 ANN. REV. NEUROSCIENCE 167 (2001).

word (e.g. "blue" printed in blue ink—congruent trials) or in a different color (e.g. "blue" written in red ink—incongruent trials). Accurately naming the color of the word during incongruent trials requires overriding the interference provided by the semantic content of the word.

Set shifting and task switching tasks index another flexibility process—adaptability to rule changes.[45] In the Wisconsin Card Sorting Task, for instance, a subject may be told to sort cards on the basis of color (a "color response set") and then told to shift to a different response set (e.g., shape). Diminished behavioral flexibility is evidenced when participants continue to make choices on the basis of a prior response set instead of using the current rules (i.e. perseverative errors).[46] In the Monsell task-switching paradigm, participants view a single digit that appears within one cell of a 2 x 2 grid. If the digit appears in one of the top row cells, participants are asked to make a judgment about the magnitude of the digit. If the digit appears in the bottom row, participants make an odd-even judgment. Higher task-switching costs—the difference in participant response time between switch trials (magnitude to odd-even or visa-versa) and repetition trials (odd-even to odd-even; magnitude to magnitude)—indicates a difficulty in flexibly updating responses when the rules that guide those responses quickly change.[47]

Finally, reversal learning tasks require an individual to use dynamic feedback to override learned stimulus-response associations.[48] For example, a participant might be asked to choose between green squares and red circles on each trial of the task. After a number of trials, the participant

---

45. Cools & D'Esposito, *supra* note 43, at 120; S.R. Chamberlain et al., *Translational Approaches to Frontostriatal Dysfunction in Attention-Deficit/Hyperactivity Disorder Using a Computerized Neuropsychological Battery*, 69 BIOLOGICAL PSYCHIATRY 1192 (2010); Y. Chudasama & T.W. Robbins, *Functions of Frontostriatal Systems in Cognition: Comparative Neuropsychopharmacological Studies in Rats, Monkeys and Humans*, 73 BIOLOGICAL PSYCHOL. 19 (2006); A.A. Kehagia, G.K. Murray & T.W. Robbins, *Learning and Cognitive Flexibility: Frontostriatal Function and Monoaminergic Modulation*, 20 CURRENT OPINION NEUROBIOLOGY 199 (2010); T.W. Robbins, *Shifting and Stopping: Fronto-Striatal Substrates, Neurochemical Modulation and Clinical Implications*, 362 PHIL. TRANSACTIONS ROYAL SOC'Y LONDON SERIES B, BIOLOGICAL SCI. 917 (2007).

46. Akira Miyake & Naomi P. Friedman, *The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions*, 21 CURRENT DIRECTIONS PSYCHOL. SCI. 8 (2012).

47. S. Monsell, *Task Switching*, 7 TRENDS COGNITIVE SCI. 134 (2003); G. Samanez-Larkin et al., *A Thalamocorticostriatal Dopamine Network for Psychostimulant-enhanced Human Cognitive Flexibility*, 74 BIOLOGICAL PSYCHIATRY 99 (2013).

48. Kehagia et al., *supra* note 45, at 201; Michael J. Frank & Eric D. Claus, *Anatomy of a Decision: Striato-Orbitofrontal Interactions in Reinforcement Learning, Decision Making, and Reversal*, 113 PSYCHOL. REV. 300 (2006); Alicia Izquierdo & J. David Jentsch, *Reversal Learning as a Measure of Impulsive and Compulsive Behavior in Addictions*, 219 PSYCHOPHARMACOLOGY 607 (2012).

learns that choosing a green square wins them $1, but choosing a red circle makes them lose $1. Without warning, the response contingency reverses (i.e. green squares lose and red circles win). Adaptive responding requires subjects to use the feedback they get on each trial (i.e. the gain or loss information following each choice) to dynamically update their choice behavior. Continuing to choose green squares after the contingency reversal would constitute a perseverative error, indicating that the subject has difficulty using punishment information (loss) to flexibly alter habitual patterns of behavior.

4. Related Constructs

Two additional constructs with conceptual links to self-control warrant mention because of their likely relevance to the law. The first, *sensation-seeking*, describes a desire for excitement or biological arousal that makes it more difficult to suppress risky or dangerous behaviors.[49] It is typically assessed via self-report questionnaires, such as the Zuckerman Sensation-Seeking Scale.[50] High levels of both impulsivity (low self-control) and sensation-seeking are found in substance abusers,[51] though each may predispose distinct aspects of drug addiction. For example, sensation-seeking may selectively increase risk for initiating substance use due to the need to seek intense sensations and willingness to tolerate risks to attain them, while impulsivity may selectively increase risk for developing substance dependence.[52]

Another relevant construct pertains to *mental representations* at the time of a decision. According to Fuzzy Trace Theory (FTT) people form two types of mental representations about an event: verbatim and gist

---

49. Marvin Zuckerman & Katherine Link, *Construct Validity for the Sensation-Seeking Scale*, 32 J. CONSULTING CLINICAL PSYCHOL. 420 (1968); Marvin Zuckerman, *Impulsive Sensation Seeking and Its Behavioral, Psychophysiological and Biochemical Correlates*, 28 NEUROPSYCHOBIOLOGY 30 (1993); David H. Zald et al., *Midbrain Dopamine Receptor Availability Is Inversely Associated with Novelty-Seeking Traits in Humans*, 28 J. NEUROSCIENCE 14372 (2008); Valerie Reyna et al., *Neurobiological and Memory Models of Risky Decision Making in Adolescents Versus Young Adults*, J. EXPERIMENTAL PSYCHOL. 1125 (2011); Britain Mills, Valerie F. Reyna & Steven Estrada, *Explaining Contradictory Relations Between Risk Perception and Risk Taking*, 19 PSYCHOL. SCI. 429 (2008).

50. Marvin Zuckerman, *Sensation Seeking and Behavior Disorders*, 45 ARCHIVES GEN. PSYCHIATRY 502 (1988).

51. Marci R. Mitchell & Mac N. Potenza, *Addictions and Personality Traits: Impulsivity and Related Constructs*, 1 PERSONALITY & IMPULSE CONTROL DISORDERS 1 (2014).

52. Karen D. Ersche et al., *Drug Addiction Endophenotypes: Impulsive Versus Sensation-Seeking Personality Traits*, 68 BIOLOGICAL PSYCHIATRY 770 (2010).

traces.[53] Gist traces are "fuzzy" representations of an event because they focus on its bottom-line meaning, whereas verbatim traces are detailed representations of an event. The type of representation used at the time of decision can profoundly affect decision-making. Adults and mature adolescents rely more on gist representations of options (e.g., unprotected sex is risky) rather than literal or verbatim representations (e.g., the risk of pregnancy is 1 in 12, of chlamydia is 1 in 50, etc.). This focus on gist highlights core values and reduces susceptibility to interference from distractions, thereby decreasing maladaptive decision-making.[54] More generally, changes in mental representations enhance self-control without requiring greater willpower.[55]

B. *Caveats*

While cognitive psychology and neuroscience research have made great strides in identifying self-control related cognitive processes and the neural circuits that serve them, several significant caveats are worth noting. First, our understanding of the latent cognitive architecture of self-control is still impoverished. Neuroscientists and cognitive scientists use experimental tasks that they think map on to the processes described above, but they are still unclear about the selectivity of those mappings, or indeed, whether the cognitive processes we suspect the tasks measure truly "carve nature at the joints." This caveat is not specific to the construct of self-control. The lack of a valid taxonomy of cognitive processes has been identified as a field-wide issue.[56] We do not yet have a consensus

---

53. Valerie F. Reyna & Charles J. Brainerd, *Dual Processes in Decision Making and Developmental Neuroscience: A Fuzzy-Trace Model*, 31 DEV. REV. 180 (2011).

54. Reyna et al., *supra* note 49, at 1132–34; Mills et al., *supra* note 49, at 431–32. These articles describe research suggesting that a focus on gist cues highlights core values because core values are stored in a gist format in long-term memory. A focus on gist also reduces susceptibility to interference from distractions because research has shown that verbatim representations of information (aka verbatim memories) are more fragile (easier to disrupt) than gist. Both of these mechanisms—cuing of core values and the robustness of gist representations (relative to verbatim representations) —are theorized to contribute to decreasing maladaptive decision-making, which has been observed. *See also* Kentaro Fujita & H. Anna Han, *Moving Beyond Deliberative Control of Impulses: The Effects of Construal Levels on Evaluative Associations in Self-control*, 20 PSYCHOLOGICAL SCIENCE 799,(2009); Valerie F. Reyna & Britain A. Mills, *Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-Trace Theory*, 143 J. EXPERIMENTAL PSYCHOL. GEN. 1627 (2014).

55. Eran Magen, Carol S. Dweck & James J. Gross, *The Hidden-Zero Effect: Representing a Single Choice as an Extended Sequence Reduces Impulsive Choice*, 19 PSYCHOL. SCI. 648 (2008); Reyna & Mills, *supra* note 54, at 1646.

56 Poldrack et al., *supra* note 33, at 18.

understanding of the discrete processes that comprise cognitive constructs. Also unclear is how such processes map onto specific experimental tasks.[57] Most tasks likely index multiple cognitive processes,[58] and performance on different tasks that purport to measure the same process often correlate only weakly,[59] reflecting the ambiguity of even well-studied cognitive constructs.

Furthermore, even discrete processes (in the sense that they rely on distinct neural circuits and are behaviorally dissociable) likely interact, both with each other and with affective and motivational context. For example, it is not clear whether an individual's performance on a simple motor response inhibition task that uses neutral visual cues (e.g., consonants and vowels) is correlated with his or her performance on a second response inhibition task that requires suppressing responses to highly valued stimuli (e.g. monetary, drug-related, sexual cues).[60] Indeed, Casey and her colleagues showed that while performance on a simple Go/NoGo task did not correlate with intertemporal choice behavior, performance on an emotional Go/NoGo task that used affectively salient stimuli Go/NoGo cues did.[61]

Two final methodological problems are particularly salient in the legal setting. The first is that lab-based measures of self-control of the type we have described may not predict self-control in the real world. Although some lab-based measures of self-control have been shown to predict some aspects of real-world self-control,[62] there is still much debate about the conditions under which this assumption is valid, and its applicability to criminal behavior is far from clear.

---

57. David Badre, *Defining an Ontology Of Cognitive Control Requires Attention to Component Interactions*, 3 TOPICS COGNITIVE SCI. 217 (2011).

58. Poldrack et al., *supra* note 33, at 18.

59. Michael J. Kane et al., *Working Memory, Attention Control, and the N-Back Task: A Question of Construct Validity*, 33 J. EXPERIMENTAL PSYCHOL. 615 (2007).

60. Erika Pike et al., *Drug-Related Stimuli Impair Inhibitory Control in Cocaine Abusers*, 133 DRUG & ALCOHOL DEPENDENCE 768 (2013).

61. B.J. Casey et al., *Behavioral and Neural Correlates of Delay of Gratification 40 Years Later*, 108 PROC. NAT'L ACAD. SCI. U.S. 14998 (2011).

62. Angela L. Duckworth, E. Tsukayama & T.A. Kirby, *Is It Really Self-Control? Examining the Predictive Power of the Delay of Gratification Task*, 39 PERSONALITY & SOC. PSYCHOL. BULL. 843 (2013); Angela L. Duckworth, P.D. Quinn & E. Tsukayama, *What Leaves Behind: The Roles of IQ and Self-Control in Predicting Standardized Achievement Test Scores and Report Card Grades*, 104 J. EDUC. PSYCHOL. 439 (2012); Reyna et al., *supra* note 49, at 1125–1140 (both behavioral intentions and self-reported risk taking have been validated as indices of real-world risk taking, i.e. they correlate with real-world risk taking); D. Goldenberg et al., *Neural Mechanisms of Impulse Control in Sexually Risky Adolescents*, 6 DEVELOP. COGNITIVE NEUROSCIENCE 23 (2013); Vivian Zayas, Walter Mischel & Gayathri Pandey, *Mind, Brain, and Delay of Gratification*, *in* THE NEUROSCIENCE OF RISKY DECISION MAKING (V. Reyna & V. Zayas eds., 2014).

Also of concern is that neuroscience data, like most scientific data, is group-based. Courts usually assume that scientific methods used to make general inferences about natural phenomena are valid for making specific inferences about individual defendants. For example, if *group-level* brain imaging, lesion, or brain stimulation studies indicate that brain region X is causally involved in a given self-control process, it is presumed that measuring an individual's fMRI signal in region X provides an objective marker of *that individual's* capacity for self-control. However, the translation of general scientific findings into information that is useful in an individual case is far from straightforward,[63] a complicated inferential process that has been called the "G2i" problem.[64]

The G2i problem is especially acute when cognitive/neuro-scientific data are used to make determinations about individual self-control capacity. Valid classification is predicated on knowledge of the sensitivity, specificity, and predictive value of one's measures; however, these basic data are largely lacking. Further, while the psychometric properties of many cognitive tasks are reasonably well established, we know little about the test-retest reliability or split-half reliability of most fMRI-based tests of self-control. Finally, as noted above, there exist no large, demographically representative datasets of the kind required for standardization of measures. This means that, as of now, even the best experimental *tasks* are not useful as *tests* for determining capacity in any individual defendant. Without these large-scale population norms on performance and brain activity for each legally relevant task, even measures with high reliability and validity will be unsuitable for individual-level inference of the kind that is required in courtroom settings.[65]

V. SELF-CONTROL IN THE COURTROOM: EVIDENTIARY CONSIDERATIONS

With this brief background on the law and science of self-control in mind, imagine a man named John is charged with murder and wants to argue that he has "a substantial inability to conform [his] behavior to the

63. Michael T. Treadway & Joshua W. Buckholtz, *On the Use and Misuse of Genomic and Neuroimaging Science in Forensic Psychiatry: Current Roles and Future Directions*, 20 CHILD & ADOLESCENT PSYCHIATRIC CLINICS N. AM. 533 (2011).

64. David L. Faigman, John Monahan & Christopher C. Slobogin, *Group to Individual (G2i) Inference in Scientific Expert Testimony*, 81 U. CHI. L. REV. 417, 420 (2014) ("This gap between conventional scientific practice and ordinary trial practice involves the challenge of reasoning from group data to decisions about individuals (an analytical process that we designate as 'G2i').").

65. Casey et al., *supra* note 61, at 150001; Duckworth et al., *Is It Really*, *supra* note 62, at 854–55.

requirements of the law," language found in both insanity doctrine and sentencing law. Assume that the methodological and conceptual problems just described are resolved (a big assumption, but one that is necessary to get to the issues we want to discuss). Assume further that John performs "poorly" on tests measuring response inhibition, action cancellation, delay discounting, and response reversal and that we can be sure John was not malingering during the testing process (another big assumption). Are the results of those tests relevant to the legal claim? The answer to that question requires answers to three sub-questions: the *type* of impairment the law considers salient, the *degree* of impairment the law considers salient, and the extent to which the law is willing to adopt a true lingua franca.

A. *Type of Impairment*

As of now, the first sub-question—the type of neurocognitive impairment that the law considers relevant—is impossible to answer, because the law's language is incompatible with cognitive and neuro-scientific measurement. Legal language referring to "substantial incapacity to conform behavior," "irresistible impulses," or "extreme mental or emotional stress" does not map well onto any of the self-control processes we have described. If legal policymakers want to take advantage of cognitive and neuro-scientific knowledge they will need to engage in hard thinking about whether difficulty in inhibiting one's responses, choosing long-range goals over short ones or any of the other self-control processes detailed above should affect legal liability.

Unfortunately, formal legal doctrine defining criminal liability does not track even the rough distinctions that moral philosophers and social psychologists have developed.[66] The relevant case law on the volitional prong of the insanity defense—the basis of the claim John is making—usually focuses on the spontaneity of the conduct, with an emphasis on the degree of planning evidenced by the defendant or the rationality of the conduct, sometimes captured through the conceit of asking whether the act would have occurred had a police officer been standing nearby.[67] At the

66. For a general discussion of this point, *see* Herbert Fingarette & Ann Fingarette Hasse, MENTAL DISABILITIES AND CRIMINAL RESPONSIBILITY 6 (1979) (arguing that the focus of non-responsibility should be "irrational condition of mind in committing [an] offense"); Hollander-Blumoff, *supra* note 17, at 505 (arguing that "psychological research on self-control seriously indicates that criminal law may vastly underdescribe the scope of situations in which an individual lacks the ability to control her actions."); Michael L. Corrado, *The Case for a Purely Volitional Insanity Defense*, 42 TEX. TECH L. REV. 481, 482–83 (2009) ("Far from being something that should be excluded from an insanity test, the volitional prong should be the only prong.").

67. MELTON et al., *supra* note 7, at 216.

same time, none other than Warren Burger, before he became the Chief Justice of the United States Supreme Court, stated that the irresistible impulse notion "has always been a misleading concept because it has connotations of some sudden outburst of impulse and completely overlooks the fact that people do a lot of weird and strange and unlawful things as a result of not just sudden impulse but long brooding and disturbed emotional makeup."[68] Thus, current legal pronouncements about the meaning of self-control in the insanity context are at best confusing. Similarly, case law defining premeditation speaks both of the need to show deliberation and "cool reflection" at the same time it insists that an act can be premeditated even if the decision to engage in it is instantaneous.[69] Self-control in the provocation context, if defined at all, is usually simply described as "extreme mental or emotional stress"[70] or an act of "sudden passion."[71]

Formal legal pronouncements about the relevance of self-control in the sentencing context are somewhat more useful. For instance, in *Graham v. Florida* the Supreme Court supported its decision that mandatory life-without-parole sentences for juvenile offenders are unconstitutional by noting that juveniles have "[d]ifficulty in weighing long-term consequences" and thus have a "corresponding impulsiveness."[72] In concluding that people with intellectual disability should not be eligible for the death penalty in *Atkins v. Virginia*, the Court spoke of these individuals' "diminished capacities to understand and process information, . . . to abstract from mistakes and learn from experience, . . . to control impulses."[73] However, this type of language is as specific as courts and statutes get about the relevance of self-control. Even when citing scientific findings related to self-control, the law's governing doctrine remains vague.

B. *Degree of Impairment*

Even if some or all of neuroscientists' findings on self-control are considered legally pertinent, however, a second determination will have to be made about *how much* impairment in the relevant self-control domains is required to affect the legal determination. The assumption above was that John performed poorly on his self-control tasks, but the word "poorly" was not defined. If John's case were tried today, neurologists or psychiatrists

68. Wion v. United States, 325 F.2d 420, 426 n.7 (10th Cir. 1963) (quoting Warren Burger).

69. Commonwealth v. Coleman, 747 N.E.2d 666, 669 (Mass. 2001).

70 State v. Haili, 79 P.3d 1263, 1267 (Haw. 2003).

71. States v. Smith, 949 S.W.2d 947, 950 (Mo. Ct. App. 1997).

72. Graham v. Florida, 560 U.S. 48, 78 (2011).

73. Atkins v. Virginia, 536 U.S. 304, 318 (2002).

testifying on his behalf would likely make statements to the effect that, based on their experience and their evaluation of John, his capacity to appreciate the wrongfulness of his conduct was "significantly impaired" or, alternatively, his capacity to control his behavior was, "within a reasonable degree of psychological certainty," significantly compromised. Unfortunately, this type of testimony is close to meaningless. Whether impairment is "significant" cannot be ascertained without data that permits a comparison of John's performance on a given self-control task to the distribution of performance on that task across a larger population of individuals. At present, we lack large, demographically representative datasets of the kind required to estimate population norms and standardize measures in a way that would be needed for compelling individual-level inference.

If such data were available then a lingua franca in the self-control setting might identify "cut points," or the number of standard deviations below which a person is considered seriously impaired for legal purposes. There is precedent for this approach. In *Castaneda v. Partida* the Supreme Court had to decide whether, in a jurisdiction that was 79.1% Latino/a, the fact that only 339 out of 870 jury pool members were Latino/a constituted a prima facie violation of the equal protection clause.[74] To answer that question, the Court resorted to statistical analysis, stating that "[a]s a general rule for such large samples, if the difference between the expected value and the observed number is greater than two or three standard deviations, then the hypothesis that the jury drawing was random would be suspect to a social scientist."[75] The Court noted that the standard deviation on the facts of *Castaneda* was 12, and went on to hold that a prima facie case of discrimination had been made out.

A second Supreme Court case that adopted scientific concepts, in a context more closely related to the self-control inquiry, is *Hall v. Florida*, decided in 2014.[76] Twelve years earlier, in *Atkins v. Virginia*, the Court had held that executing people with intellectual disability is unconstitutional even when they have intentionally killed another, in large part because, as noted above, it concluded that such people are less likely to be able to "process the information of the possibility of execution as a penalty and, as a result, control their conduct based upon that information."[77] In *Hall*, the Court reaffirmed that holding and held that, in defining intellectual disability for purposes of *Atkins'* exemption, the states must adhere to the American Psychiatric Association's formulation of that condition, which,

74. Castaneda v. Partida, 430 U.S. 482 (1977).
75. *Id.* at 496 n. 17.
76. Hall v. Florida, 134 S. Ct. 1986 (2014).
77. *Atkins*, 536 U.S. at 320.

inter alia, provides that a person is intellectually disabled if he or she has an IQ score that is "approximately two standard deviations or more below the population mean, including a margin for measurement error (general 5% points)".[78] In other words, the Court was willing to incorporate a scientific definition into a legal culpability principle.

After *Castaneda* and *Hall*, the lingua franca in jury selection and death penalty cases speaks in terms of standard deviations. In the self-control context, the law could take a similar approach. For instance, it might establish that individuals whose capacities, as measured by the types of tests described above, are two standard deviations below the mean would be entitled to a presumption of mitigation and that those whose capacities are three standard deviations below the mean would be presumptively entitled to a complete excuse. Of course, for this type of threshold to have any impact, neuroscientists would have to be able to quantify—at the level of the individual—the self-control capacities relevant to the law. For the reasons suggested above, that goal will be much harder to reach than calculating population differentials in jury selection cases or assessing IQ in death penalty cases. Further, the usefulness of the standard deviation as a measure of variance depends on sample size, the normality of the distribution, and the psychometric properties of the test. But if all of this can be worked out, policymakers and decision-makers could benefit immensely. Experts could identify relatively precisely against a known baseline how impaired an individual is in a given self-control capacity.

C. *Potential Objections*

Would the law be willing to adopt such a lingua franca? Although *Casteneda* and *Hall* demonstrate that courts are sometimes willing to move in that direction, lawyers have often resisted "trial by mathematics," especially in connection with legal issues that go to substantive liability, for at least three reasons. First, there is a justifiable distrust of numbers. In the self-control context, until cognitive science and neuroscience can demonstrate they possess the empirical foundation needed for compelling individual inference, the law should be hesitant to invite them into the courtroom.

---

78. AMERICAN PSYCHIATRIC ASSOCIATION, DIAGNOSTIC AND STATISTICAL MANUAL 37 (5th ed. 2013), quoted in Hall v. Florida, 134 U.S. at 1995. The definition also requires that the individual evidence "deficits in adaptive functioning that result in failure to meet development and socio-cultural standards for personal independence and social responsibility." *Id.* This aspect of the definition is not as well-normed, although instruments with psychometric properties have been designed to measure it. *See, e.g.*, PATTI L. HARRISON & THOMAS OAKLAND, ADAPTIVE BEHAVIOR ASSESSMENT SYSTEM (3d. ed., 2016), available at http://mhs.com/product.aspx?gr=edu&prod=abas3&id=overview.

Second, there is a valid concern that, even if the science is solid, attempting to fashion a lingua franca with it might tempt the law to compromise its normative premises. After *Atkins*, for instance, legal scholars worried that the diagnosis of intellectual disability, developed primarily for treatment purposes, did not mesh with the law's focus on blameworthiness.[79] The Supreme Court's subsequent decision in *Hall* ignored this concern, implicitly reaffirming *Atkins'* conclusion that, even if the fit between clinical and legal concepts in this context was imperfect, no one with intellectual disability was among the worst of the worst offenders who could justifiably be put to death. However, in other domains the overlap between what science can produce and what the law needs might turn out to be so minimal that any attempt at developing a lingua franca will make matters worse.

Less justifiably, the legal system is also squeamish about reducing culpability determinations to probabilities and other quantified concepts.[80] The intuition here is that one's "score" on the Go/NoGo or delay discounting task, even if clearly valid and clearly indicative of legally relevant deviations from "normal" cognitive capacity, should never be allowed to inform conclusions about a particular level of culpability or level of punishment. The most likely explanation for this adamant position is the concern that giving significant weight to scientific data—and to scientists—has the effect of reducing the law's power to control the scope of the inquiry.[81]

Precisely this sort of tension over the quantification of legal concepts arose in *Hall*. The dissent in that case criticized the majority's definition of intellectual disability on the ground that it substituted the consensus of "professional societies" for the consensus of the general public and its representatives.[82] In particular, the dissent did not agree with the majority's conclusion that the states must take into account the American

---

79. *See* Lois A. Weithorn, *Conceptual Hurdles in the Application of Atkins v. Virginia*, 59 HASTINGS L.J. 1203, 1223 (2008) (Tests measuring intellectual disability "were not developed for the purpose of distinguishing between capital offenders whose deficits in intellectual functioning rend them ineligible for the death penalty from those capital offenders without such deficits").

80. *See* Laurence H. Tribe, *Trial By Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1331 (1971) ("in at least some contexts, permitting *any* use of certain mathematical methods entails a sufficiently high risk of misuse, or a risk of misuse sufficiently costly to avoid, that it would be irrational not to take such misuse into account when deciding whether to permit the methods to be employed at all.").

81. *Cf.* David Faigman, *Normative Constitutional Fact-finding: Exploring the Empirical Component of Constitutional Interpretation*, 139 U. PA. L. REV. 541, 581–88 (1991) ("[u]ltimately, persistent misapplication of empirical data or, in this case, the ignoring of such data undermines the Court's legitimacy.").

82. *Hall*, 134 S. Ct. at 2002.

Psychiatric Association's two standard deviation referent for IQ scores (which amounts to an IQ of 70) or the standard error of measurement, also endorsed by the APA, in deciding whether an offender has such a score. But six members of the Court were nonetheless willing to adopt the APA's position, on the ground that it captured the normative premise of *Atkins* and more accurately identified those who met it. After *Hall*, while juries and judges will still be the arbiters of whether a person is intellectually disabled, their inquiry will be governed by scientific concepts. Whatever one thinks about the holding in *Hall*, it is a prime example of a lingua franca approach.

While the law's concerns about "scientizing" culpability and other normative concepts must be taken seriously, they should not prevent the endeavor from going forward in a cautious manner. Ideally, an iterative dialogue between the two disciplines would identify common ground. As in *Hall*, the law could then formulate that ground in terms that scientists can understand and address. The final section describes this iterative process in more detail.

VI. THE PATH FORWARD

The relationship between neuroscience and criminal law is still in its infancy. If neuroscience is to be maximally useful to law, it is imperative that both disciplines work *proactively* to develop a coherent framework for dealing with neuro-scientific evidence, rather than reactively responding in an ad hoc manner each time such evidence is considered in the courts. Here, we focus on self-control, a key zone of engagement. We propose that, to the extent the degree of self-control is considered legally relevant, the law adopt a lingua franca that relies on cognitive and neuro-scientific definitions in articulating the relevant concepts.

What might this lingua franca look like? While we believe that the details must reflect the outcome of an iterative, multilateral process that is open to all relevant stakeholders (e.g. legal scholars, judges, litigators, cognitive scientists, neuroscientists, and forensic clinicians), a few concrete goals can be outlined, represented in Figure 1 at the end of this article. First, a lingua franca of self-control should facilitate objective classification of legal standards according to scientifically meaningful criteria. In practice, this process would begin with an exhaustive survey of case law, legal opinions, state penal codes, and law review articles, with the goal of generating a definitive list of self-control standards applied in courtrooms.

The next step would involve identifying sets of experimental paradigms that putatively operationalize and quantify the capacities described by each selected legal standard. Each standard could be mapped to 2 or 3 candidate experimental paradigms. For example: "ability to

control complex bodily function" is one standard; the cognitive paradigms that might map most closely are motor response inhibition tasks, such as the stop-signal and go/no-go paradigms. By contrast, another standard—"the ability to contemplate consequences of one's actions"—might map more closely to related action-outcome learning and cost-benefit decision-making tasks. The outcome of this stage would be a hypothetical task-standard map.

The validity of such maps could be assessed by surveying state and federal judges (perhaps through the National Center for State Courts, the Administrative Office of U.S. Courts, and the Federal Judges Association) to measure the degree of agreement among jurists about the relationship between standards and measures. Behavioral scientists would translate methodological details into comprehensible text for jurists, who would in turn indicate the degree to which they believe each paradigm measures the mapped legal standard. For example, one survey question could describe a standard "Go/NoGo" paradigm and associated scientific evidence; respondents would then be asked to indicate which of several legal self-control standards is most accurately assessed by that paradigm.

Crucially, such a lingua franca must demarcate specific legal standards for which there is consistent agreement that no viable scientific operationalization exists, or for which the law indicates it neither desires nor requires scientific evidence to adjudicate. The legitimacy of the lingua franca enterprise rests largely on this third step. The law is always free to decide that it does not need or want science to adjudicate legal standards. If jurists do decide to make decisions based on their own moral intuitions or the normative preferences and precedents of their field, however, they should not then use scientific evidence as an "empirical fig-leaf" to prop up the intuitions and preferences that are in fact driving their judgments.

In sum, good faith, as opposed to merely instrumental, use of science in the courts requires legal standards to be framed (or re-framed, as appropriate) in a manner that makes them accessible to valid scientific operationalization and inference. That does not require that scientific terms replace the legal ones, as in *Hall*. The law can always retain its statutory and judicial language defining the ultimate self-control issues. But if it wants useful and valid scientific information, it should also develop sub-tests or presumptions that incorporate empirically verifiable constructs. In other words, if the law wishes, or intends to permit, the use of cognitive and neuro-scientific evidence to make inferences about legally germane aspects of mind and brain, legal policymakers must work with scientists to ensure that their constructs map on to scientific valid data.

VII. CONCLUSION

The underlying premise of this article is that if the law wants the full benefits of neuro-scientific knowledge, it needs to develop a lingua franca based on neuro-scientific concepts. In doing so, the law should be mindful of the limits of scientific inference (particularly as it pertains to legally relevant individual-level assessment) and alert to the fact that, despite semantic similarities, scientific constructs often do not track with its normative precepts. But the law should not resist a neuro-scientific lingua franca simply because it conflicts with legal sensitivities or surrenders too much power to scientists.

The effort in this article to reconcile legal standards for self-control with current neuro-scientific insights is unavoidably preliminary in nature. Such efforts must begin now because of the accelerating pace with which neuroscience evidence is being introduced in the courts. Until such efforts bear fruit, law and neuroscience will continue its high-stakes game of charades, each making stabbing guesses at what the other means, while citizens' lives and freedom hang in the balance.

Figure 1: Experimental Tasks, Cognitive Processes, and Legal Standards Related to Self-Control

The boxes here (at the top) highlight several legal standards that might benefit from an evaluation of individual cognitive capacities related to self-control in some or all of the three putative domains of self-control - action, choice, and flexibility – represented by the pyramids. Domain-specific cognitive processes for each domain are identified underneath the pyramids. Experimental task paradigms are grouped here according to the domain of self-control that they access (at the bottom), as well as the processes within each domain that each paradigm measures (immediately below the triangles). Note that there is significant debate about the latent cognitive structure of self-control, and particularly about specific task-process mappings for many of these paradigms. We do not intend this figure to be a definitive characterization of what is a dynamic and evolving area of science.