Duke Law Journal Online

VOLUME 71 OCTOBER 2021

NEUROSCIENCE AND THE MODEL PENAL CODE'S *MENS REA* CATEGORIES

ANDREAS KUERSTEN & JOHN D. MEDAGLIA†

ABSTRACT

This Essay addresses recent research and commentary regarding the potential contributions of cognitive neuroscience to law. For the first time, cognitive neuroscience methods have been brought to bear on the Model Penal Code's (MPC's) culpable–mental state categories through a neuroimaging study seeking to identify the neural correlates of knowledge and recklessness. Subsequently, this study has been presented as a paradigm for utilizing cognitive neuroscience to answer important legal questions. However, the original experiment appears to suffer serious experimental-design and conceptual limitations, belying subsequent advocacy for the legal utility of cognitive neuroscience. This Essay methodically details these limitations and argues that the original study does not seem to have actually elicited knowledge or recklessness in subjects or, even if it did, did not elicit them in discrete enough fashion to permit identification of the mental states' neural correlates. The Essay also contends, more broadly, that cognitive neuroscience appears inapt for investigating the propriety of the MPC's mens rea *delineations since these are articulated in purely psychological-behavioral terms: mental states are the only requisites. Only psychological-behavioral manifestations provide base evidence of mental states' existences. And psychological-behavioral research, not cognitive neuroscience, is the most direct way to investigate the practical, moral, and legal appropriateness of the MPC's mental states by illuminating how individuals experience them, identify them in*

† Andreas Kuersten is Visiting Researcher with the Neuroethics Studies Program at Georgetown University Medical Center's Center for Clinical Bioethics; John D. Medaglia is Assistant Professor of Psychology and Neurology at Drexel University and Adjunct Assistant Professor of Neurology at the University of Pennsylvania. The authors would like to thank Teneille Brown, Sarah E. Coco, Geoffrey E. Curfman, David DeMatteo, Kirk Heilbrun, and Stephen J. Morse for helpful exchanges and comments at various stages in the production of this Essay. The authors would also like to thank the Duke Law Journal editors for their efforts and suggestions, which were markedly beneficial. A special thanks is reserved for Catherine Kuersten for comprehensively reviewing this piece. Any mistakes or omissions are solely those of the authors.

others, or employ them to dispense blame and punishment. Ultimately, recent cognitive neuroscience research does not appear to reveal anything of legal significance regarding the MPC. And, more broadly and contrary to recent assertions, cognitive neuroscience has substantial limitations when it comes to producing legally relevant information. Going forward, psychological-behavioral research should be given primacy in cognitive science investigations of MPC concepts. Cognitive neuroscience studies, on the other hand, should be treated with exceptional skepticism.

TABLE OF CONTENTS

INTRODUCTION

As the Model Penal Code (MPC) approaches its sixth decennial, the legal innovations it embodies continue to reverberate through criminal law doctrine and scholarship alike.[1] And perhaps no contribution is as substantial as the MPC's authors recognizing four culpable mental states: purpose, knowledge, recklessness, and negligence.[2] Given the prominence and influence of this delineation,

1. *See* MARKUS D. DUBBER, AN INTRODUCTION TO THE MODEL PENAL CODE 1–7 (2nd ed., 2015) (remarking on the continued, expansive, and unparalleled influence of the MPC on criminal law); Paul H. Robinson & Markus D. Dubber, *The American Model Penal Code: A Brief Overview*, 10 NEW CRIM. L. REV. 319, 340 (2007) ("For almost half a century, the [MPC] has been the dominant force in American criminal code reform and a catalyst for American criminal law scholarship.").

2. MODEL PENAL CODE § 2.02(2); *see* Luis E. Chiesa, Mens Rea *in Comparative Perspective*, 102 MARQ. L. REV. 575, 579 (2018) ("The most influential provision of the MPC is [the] section . . . defin[ing] subjective offense elements," which defines culpable mental states);

its frequent evaluation and critique over decades is unsurprising.[3] But despite this extensive engagement, there remain avenues of study that are relatively unworn or, indeed, entirely untraversed. One such avenue is a methodology employed by a team led by psychologist Iris Vilares (Vilares and colleagues) for the first time roughly 55 years after the MPC's official release: cognitive neuroscience utilizing functional magnetic resonance imaging (fMRI).[4]

Cognitive neuroscience is the study of how the brain gives rise to the mind.[5] For its part, fMRI, along with attendant technology and methods, is a powerful tool for studying brain function and, pertinently, identifying the neural correlates of cognitive activity.[6] Put simply, fMRI indirectly and non-invasively measures brain function by indirectly measuring blood flow: neuronal activity is correlated with a delayed influx of oxygenated blood—e.g., as neuronal activity in a brain region increases, the amount of oxygenated blood in that region increases as well—and fMRI measures a signal (the blood oxygen level–dependent (BOLD) signal) correlated with the amount of

---

Robinson & Dubber, *supra* note 1, at 335 ("[T]he [MPC]'s definitions of these four mental states may be the code's most important contribution to American criminal law reform."). "Mental state" is a contested psychological concept, though it is generally used to describe states of thinking and feeling. E.J. LOWE, AN INTRODUCTION TO THE PHILOSOPHY OF MIND 39 (2000). A notable, though certainly not controlling, recent definition of mental state is the following:

> A transient cognitive or emotional state of the organism that can be described in terms of its contents (what the state is 'about') and the relation that the subject bears to the contents (for example, perceiving, believing, fearing, imagining or remembering).

Kalina Christoff, Zachary C. Irving, Kieran C.R. Fox, R. Nathan Spreng & Jessica R. Andrews-Hanna, *Mind Wandering as Spontaneous Thought: A Dynamic Framework*, 17 NATURE REV. 718, 719 (2016). In the MPC, relevant mental states are defined in terms of the intentionality or awareness, or lack thereof, that they manifest towards the elements of delineated criminal offenses. Paul H. Robinson, *Mens Rea*, *in* ENCYCLOPEDIA OF CRIME & JUSTICE 995, 998–1001 (Joshua Dressler ed., 2nd ed., 2002); *see infra* note 44 and accompanying text.

3. *See infra* notes 55–88 and accompanying text.

4. Iris Vilares, Michael J. Wesley, Woo-Young Ahn, Richard J. Bonnie, Morris Hoffman, Owen D. Jones, Stephen J. Morse, Gideon Yaffe, Terry Lohrenz & P. Read Montague, *Predicting the Knowledge-Recklessness Distinction in the Human Brain*, 114 PROC. NAT'L ACAD. SCI. 3222 (2017).

5. STEPHEN M. KOSSLYN & OLIVIER KOENIG, WET MIND: THE NEW COGNITIVE NEUROSCIENCE 3–4 (1992); *see generally* Michael I. Posner & Gregory J. DiGirolamo, *Cognitive Neuroscience: Origins and Promise*, 126 PSYCH. BULL. 873 (2000) (describing the field of cognitive neuroscience and its history). Cognitive neuroscientists "typically start with the assumption that the [physicalist] view of the mind . . . is correct (that the mind is a product of the brain), and then try to understand how the brain makes the mind possible." JOSEPH LEDOUX, SYNAPTIC SELF: HOW OUR BRAINS BECOME WHO WE ARE 18 (2002).

6. *See* PETER A. BANDETTINI, fMRI 23–50 (2020) (presenting fMRI technology and techniques and their history). fMRI is considered "the preeminent method of functional neuroimaging." Geoffrey K. Aguirre, *Functional Neuroimaging: Technical, Logical, and Social Perspectives*, 44 HASTINGS CTR. REP. S8, S8 (2014).

oxygenated blood in brain regions over time.[7] Researchers can discern brain regions that are relatively more or less active during a cognitive task by comparing the BOLD signal in the regions during the target task to that during alternative tasks or baseline states—i.e., by analyzing the BOLD contrast.[8]

Vilares and colleagues employ cognitive neuroscience and fMRI in an attempt to build on a body of psychological-behavioral research—i.e., research examining mental states and behavior through subjects' behavioral responses to tasks—investigating the legitimacy of the MPC's culpable mental states.[9] Authors of the studies comprising this body of literature largely interpret their findings as suggesting that individuals have difficulty recognizing and discriminating between the MPC's knowing and reckless mental states and, contrary to the MPC's directives, do not consistently punish those who exhibited knowledge during an offensive act more than those who exhibited recklessness.[10] These results potentially challenge the propriety of knowledge and recklessness as separate culpability determinants and, perhaps, their existences as distinct mental states.[11] Vilares and colleagues seek to address these two issues by attempting to ascertain whether knowledge and recklessness entail different brain activity.[12]

In a recent essay, *Detecting Mens Rea in the Brain*, a team led by legal scholar Owen Jones (Jones and colleagues) expounds upon and

---

7. *Id*. at S9–S10; BANDETTINI, *supra* note 6, at 26–29. According to psychiatrist Sally Satel and psychologist Scott Lilienfeld,

> [fMRI] leverages the fact that everything the brain enables us to do—feel, think, perceive, and act—is linked, or correlated, with changes in oxygen consumption and regional blood flow in the brain. When a person responds to a task, . . . specific regions of the brain are typically engaged and receive more oxygen-laden, or oxygenized, blood. The increased blood flow and the boost in oxygen associated with it are proxies for increased activation of neurons.

SALLY SATEL & SCOTT O. LILIENFELD, BRAINWASHED: THE SEDUCTIVE APPEAL OF MINDLESS NEUROSCIENCE 5 (2013). For an in-depth assessment of fMRI, see generally Richard B. Buxton, *The Physics of Functional Magnetic Resonance Imaging (fMRI)*, 76 REP. PROGRESS PHYSICS 1 (2013).

8. Teneille Brown & Emily Murphy, *Through a Scanner Darkly: Functional Neural Imaging as Evidence of a Criminal Defendant's Past Mental States*, 62 STAN. L. REV. 1119, 1139–40 (2010); *see* Adina L. Roskies, *Brain Imaging Techniques*, *in* A PRIMER ON CRIMINAL LAW AND NEUROSCIENCE 61 (Stephen J. Morse & Adina L. Roskies eds., 2013) ("Almost all neuroimaging experiments are interested in comparing data between different task conditions.").

9. Vilares et al., *supra* note 4, at 3222–23.

10. *See infra* notes 57–88 and accompanying text. Being determined to have been knowing rather than reckless can often have dramatic implications for defendants in terms of potential and actual punishment. Owen D. Jones, Read Montague & Gideon Yaffe, *Detecting Mens Rea in the Brain*, 169 U. PA. L. REV. 1, 2–3 (2021).

11. *See infra* note 86 and accompanying text.

12. Vilares et al., *supra* note 4, at 3222–23, 3227.

extols the aforementioned study, of which they are coauthors.[13] They posit that Vilares and colleagues discerned distinct physical underpinnings of knowledge and recklessness and that this provides scientific support for these classifications being actual, separate psychological states and, therefore, appropriate determinants of legal culpability.[14] They also contend that "the moral legitimacy of the [MPC]'s taxonomy of culpable mental states—which punishes those in defined mental states differently—*depends on* whether those mental states actually correspond to different brain states in the way the [MPC] categorization assumes."[15] That is, specific cognitive neuroscience experimental results are allegedly not only relevant to but also necessary for legitimizing the Code's mental state categories. However, due to experimental-design and conceptual issues with Vilares and colleagues' study, the assertions of Jones and colleagues appear inadequately supported.[16]

This Essay comprehensively evaluates the study by Vilares and colleagues and the claims of Jones and colleagues. It ultimately resolves that the study was not properly designed so as to elicit knowledge or recklessness or, even if it did, it did not measure these mental states' neural correlates. In addition, cognitive neuroscience is shown to be an inapt approach to investigating the existence or legitimacy of the MPC's culpable mental states, which are psychological-behavioral concepts and, as such, can only be fundamentally validated using psychological-behavioral research methods. Thus, contrary to the assertions of Jones and colleagues, the study by Vilares and colleagues does not appear to be a paradigm of legally-relevant cognitive neuroscience research. In fact, it is questionable whether the study produced any findings of legal significance.

To be clear, this Essay does not stand against cognitive neuroscience potentially producing legally relevant findings. But the use of and expectations for these methods must be appropriately tempered. Going forward, researchers seeking to investigate the reality and propriety of the Code's *mens rea* categories via cognitive science should proceed cautiously and in an actionable progression. Any one

13. Jones et al., *supra* note 10.

14. *Id.* at 21–22.

15. *Id.* at 30 (emphasis added).

16. Vilares and colleagues also make suspect contentions. *See* Vilares et al., *supra* note 4, at 3225–27. While their article is not the main target of this Essay, Vilares and colleagues' questionable predications are largely reiterated and elaborated on by Jones and colleagues and are therefore substantially addressed here.

experiment will likely contribute relatively little of practical, moral, or legal significance unless the following are ordinally obtained: (1) clear, reliable, and valid methods for inducing and quantifying mental states; (2) equally lucid mappings between mental states and the MPC; and (3) reliable and valid neural tests that are sensitive and specific to distinctions between mental states. Given that (1) is broadly lacking, efforts should be invested there until sufficient consensus has been reached with regard to these methods, and cognitive neuroscience studies should be treated with exceptional skepticism.

This Essay proceeds in three parts. Part I provides relevant context to Vilares and colleagues' study and Jones and colleagues' claims. It concisely articulates the history of the MPC's mental state categories, their influence on criminal law, and the current body of empirical psychological-behavioral research examining this section of the MPC. In so doing, this part illuminates how the MPC's culpability criteria were conceived of in non-scientific, psychological-behavioral terms and entail purely psychological-behavioral requisites. As a result, psychological-behavioral research, not neuroscience, has thus far advanced legally relevant understandings of the MPC and the knowing and reckless categories in particular. Part II outlines the experiment conducted by Vilares and colleagues. The cognitive neuroscience methods and results are presented in detail so as to facilitate robust engagement. This part concludes with the study's alleged implications as put forth by Jones and colleagues. Part III presents our full critique of the study by Vilares and colleagues and the claims of Jones and colleagues. First, *interpretive limitations* of the experiment are described that cabin the explanatory power of Vilares and colleagues' study but do not significantly challenge Jones and colleagues' assertions. These limitations comprise the timing of the neuroimaging conducted by Vilares and colleagues and the nature of recklessness as articulated in the MPC. Second, *experimental-design limitations* are detailed that conceivably undercut Jones and colleagues' claim that the experiment discerned disparate neural functioning correlated with knowledge or recklessness. This part posits that the experiment seems improperly designed to evoke knowledge or recklessness.[17] And even if the target mental states were elicited, subjects likely did not experience knowledge and recklessness in a discrete enough fashion for attendant neural functioning to be credibly

17. Deductions as to the mental states subjects actually experienced are frustrated by Vilares and colleagues not requiring participants to explicitly register their appreciations of experimental conditions. *See infra* notes 144–149 and accompanying text.

discerned. Finally, broader *conceptual limitations* are outlined. These entail the nature of the MPC's culpable mental states as psychological-behavioral concepts whose base actuality can only be discerned via psychological-behavioral methods. Cognitive neuroscience can only contribute to investigating the aforementioned mental states once they have been sufficiently validated by psychological-behavioral research, which has not yet occurred. By prematurely employing cognitive neuroscience to verify the MPC's knowing and reckless culpability standards, Vilares and colleagues appear to have produced legally trivial findings.

I. RELEVANT CONTEXT

It is a core tenet of Anglo-American criminal law that "*actus non facit reum nisi mens sit rea* (an act does not make one guilty unless his mind is guilty)."[18] This principle is generally implemented by formulating crimes as compounds of guilty acts (*actus rei*) and guilty minds (*mentes reae*).[19] And arguably the most influential work with regard to *mentes reae* is the MPC,[20] crafted over roughly a decade by esteemed experts from multiple disciplines and promulgated in 1962.[21] To adequately foreground critiques of cognitive neuroscience research targeting the Code's knowing and reckless culpable mental states, this section first briefly expounds upon the origins of the MPC's four *mens rea* categories. It then describes relevant reported findings of the psychological-behavioral literature addressing these mental states.

---

18. Martin R. Gardner, *The* Mens Rea *Enigma: Observations on the Role of Motive in the Criminal Law Past and Present*, 1993 UTAH L. REV. 635, 636; *accord* Francis Bowes Sayre, Mens Rea, 45 HARV. L. REV. 974, 974 (1932).

19. *See* Morissette v. United States, 342 U.S. 246, 251–52 (1952) ("Crime, as a compound concept, generally constituted only from concurrence of an evil-meaning mind with an evil-doing hand, . . . took deep and early root in American soil."). The most prominent exceptions to this formulation of crimes are strict liability offenses, which require no *mens rea*, Thomas Weigend, *Subjective Elements of Criminal Liability*, *in* THE OXFORD HANDBOOK OF CRIMINAL LAW 491 (Markus D. Dubber & Tatjana Hörnle eds., 2014), and are controversial because of this, *see, e.g.*, APPRAISING STRICT LIABILITY (A.P. Simester ed., 2005) (presenting varied analyses of strict liability).

20. *See, e.g.*, Markus Dubber, *The American Law Institute's Model Penal Code and European Criminal Law*, *in* SUBSTANTIVE CRIMINAL LAW OF THE EUROPEAN UNION 210 (André Klip ed., 2011) ("[T]he [MPC] is the closest thing to a conceptual—if not a doctrinal—backbone of American criminal law."); Alan C. Michaels, *Acceptance: The Missing Mental State*, 71 S. CAL. L. REV. 953, 958 (1998) ("The [MPC] [is] widely recognized as comprising the dominant culpability scheme in American criminal law.").

21. Robinson & Dubber, *supra* note 1, at 323–24; Herbert Wechsler, *Codification of Criminal Law in the United States: The Model Penal Code*, 68 COLUM. L. REV. 1425, 1426–27 (1968).

A. The Origins of the Model Penal Code's Culpable Mental States

Prior to the MPC, criminal law "was [in] a chaotic state . . . with regard to the mental elements of offenses."[22] Widespread but uncoordinated, desultory court treatments utilized "an abundance of [ill-defined] terms,"[23] which resulted in "variety, disparity[,] and confusion" concerning "the requisite but elusive mental element."[24] The MPC refined "the dozens of mental states that . . . emerged over the course of the common law" into four that are relatively punctiliously specified: purpose, knowledge, recklessness, and negligence.[25]

Legal scholar Herbert Wechsler, the Chief Reporter and principal drafter of the MPC, describes demarcating and defining the MPC's culpable mental states as "one of the hardest drafting problems in the framing of the Code,"[26] the resolution of which was profoundly affected by his deep knowledge and personal understanding of criminal law. At the start of Wechsler's work on the MPC, he acknowledged psychologists' criticisms of then-contemporary doctrine: namely, that it employed unsound, superficial, and outmoded psychological concepts like "deliberation," "will," "passion," and "intent" in characterizing and categorizing wrongdoers and wrongdoing.[27]

---

22. Ronald L. Gainer, *The Culpability Provisions of the Model Penal Code*, 19 RUTGERS L.J. 575, 577 (1988).

23. Wechsler, *supra* note 21, at 1436. Legal scholar Luis Chiesa remarks that

> [t]his haphazard approach led courts of various jurisdictions to invoke terms such as felonious intent, criminal intent, malice aforethought, guilty knowledge, fraudulent intent, willfulness, and scienter, to denote the guilty knowledge, or mens rea, of criminal offenses. It is unclear what each of these terms mean and how they differ from each other, if they differ at all.

Chiesa, *supra* note 2, at 578–79 (citation and internal quotation marks, ellipses, and brackets omitted); *see* Paul H. Robinson, *A Brief History of Distinctions in Criminal Culpability*, 31 HASTINGS L.J. 815, 815 (1980) (estimating that American common law included "nearly eighty miscellaneous culpability terms").

24. Morissette, 342 U.S. at 252.

25. Robinson & Dubber, *supra* note 1, at 335; *see* MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962) (defining the purposeful, knowing, reckless, and negligent mental states); Dannye Holley, *The Influence of the Model Penal Code's Culpability Provisions on State Legislatures: A Study of Lost Opportunities, Including Abolishing the Mistake of Fact Doctrine*, 27 SW. U.L. REV. 229, 230 (1997) ("The Code obliterated ill-defined, confusing common law language and concepts and replaced them with four specifically defined hierarchical levels of culpability.").

26. Herbert Wechsler, *The Model Penal Code and the Codification of American Law*, in CRIME, CRIMINOLOGY AND PUBLIC POLICY: ESSAYS IN HONOUR OF SIR LEON RADZINOWICZ 435 (Roger Hood ed., 1974).

27. Herbert Wechsler, *The Challenge of a Model Penal Code*, 65 HARV. L. REV. 1097, 1103 (1952).

However, Wechsler was relatively skeptical of the potential for science to contribute to law.[28] Whereas some eminent scholars advocated science-first approaches to crafting a model penal code—i.e., first conducting criminal law–relevant scientific research and allowing any code to "emerge as a natural product of the various detailed investigations"[29]—Wechsler explicitly rejected such notions and emphasized the judgment of legal thinkers and existing law in fashioning doctrine.[30] "[O]nly by systematic study of the penal law and its pervasive problems," he writes, "can we [legal experts] appraise the relevancy of behavior science in this field."[31] And he endorsed folk-psychological, as opposed to technical-psychological, appreciations of human thought and conduct as being most pertinent to criminal law.[32] That is, he supported delineating criminal law in terms of acts and mental states described in ordinary, non-scientific language.[33]

Though Wechsler acknowledged "pervasive problems" in then-extant penal regimes, he appreciated criminal law's accumulated wisdom.[34] This makes one aspect of his approach to jurisprudence

28. *See* Gerald Leonard, *Towards a Legal History of American Criminal Theory: Culture and Doctrine from Blackstone to the Model Penal Code*, 6 BUFF. CRIM. L. REV. 691, 812 (2002) ("Wechsler always harbored a good deal more skepticism of the scientists' powers than did many other reformers.").

29. Jerome Hall, *The Proposal to Prepare a Model Penal Code*, 6 J. LEGAL EDUC. 91, 93 (1951).

30. Wechsler, *supra* note 27, at 1132–33; *see* Herbert Wechsler, *American Criminal Law Institute: II. A Thoughtful Code of Substantive Law*, 45 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 524, 530 (1955) ("The legal standard of responsibility is not a proposition in psychiatry. It is a moral and juristic concept drawn from deep ideas of justice derived from the ancient world.").

31. Wechsler, *supra* note 27, at 1132–33. *See* Wechsler, *supra* note 30, at 530 ("We must consider first the ends that law should serve before we can determine how far science bears on their attainment.").

32. *See* Herbert Wechsler & Jerome Michael, *A Rationale of the Law of Homicide: I*, 37 COLUM. L. REV. 701, 761 n.203 (1937) (responding to a critique that criminal law's "old analysis of act and intent can stand only as an artificial legal analysis" by stating that "an analysis of criminal behavior in terms of act, knowledge, intent, motive and risk is essential for legal purposes," and "[i]n so far as such an analysis is psychological, it is not in any respect unsound," for while "[i]t does not provide a complete psychology[,] the problem is not solely a psychological one") (internal quotation marks and citation omitted).

33. *See* MICHAEL S. MOORE, MECHANICAL CHOICES: THE RESPONSIBILITY OF THE HUMAN MACHINE 12–13 (2020) ("The folk psychology . . . is the psychology of rational agency in terms of which we describe, explain, and evaluate our own and others' behavior in daily life."); MICHAEL S. PARDO & DENNIS PATTERSON, MINDS, BRAINS, AND LAW: THE CONCEPTUAL FOUNDATIONS OF LAW AND NEUROSCIENCE, at xviii (2013) ("The expression 'folk psychology' refers to our common psychological/mental concepts and our ordinary use of words expressing the[m].").

34. *See* Wechsler, *supra* note 30, at 525–26 ("[W]e inherited a system at the start and here, as elsewhere, it was more congenial upon the whole to build on that foundation [of prevailing doctrine] than to start anew.").

particularly salient in contemplating the MPC's culpability criteria: his emphasis on "procedural particularity and institutional competence over substantive doctrine."[35] Legal scholar David Wolitz cogently argues that, in crafting the MPC, Wechsler did not endeavor to "remak[e] the law from a blank slate according to any single overarching normative theory."[36] Rather, he sought to glean and refine core values, principles, and purposes from existing regimes in order to "match means (doctrines) and ends (values) as effectively as possible" in a concise, coherent form readily implementable, in whole or in part, by legislatures.[37] And Wechsler viewed the primary purpose of criminal law as preventing certain harms.[38] To serve this purpose, he supported deterrence- (both special and general), incapacitation-, and treatment-focused measures.[39] (He was quite critical of retributivist framings of criminal law.[40]) These aims fundamentally influenced Wechsler's consideration of mental states with regard to criminality:

---

35. David Wolitz, *Herbert Wechsler, Legal Process, and the Jurisprudential Roots of the Model Penal Code*, 51 TULSA L. REV. 633, 634 (2016). "Wechsler was one of the leading exponents of the legal process school that dominated academic law in the 1950s and 1960s," Jonathan Simon, *Wechsler's Century and Ours: Reforming Criminal Law in a Time of Shifting Rationalities of Government*, 7 BUFF. CRIM. L. REV. 247, 247 (2003), which emphasized effective procedure and institutions over normative considerations. Some consider the MPC "the great legislative achievement of the 'legal process' school of thought in American law." Harold Edgar, *Herbert Wechsler and the Criminal Law: A Brief Tribute*, 100 COLUM. L. REV. 1347, 1355 (2000). For an in-depth treatment of legal process thinking, see generally Charles L. Barzun, *The Forgotten Foundations of Hart and Sacks*, 99 VA. L. REV. 1 (2013).

36. Wolitz, *supra* note 35, at 634.

37. *Id*. at 635; *see* Markus D. Dubber, *Penal Panopticon: The Idea of a Modern Model Penal Code*, 4 BUFF. CRIM. L. REV. 53, 76–77 (2000) ("Ever the pragmatic practitioner of Legal Process, Wechsler viewed penal law as a matter of instrumental policy."); Wechsler, *supra* note 27, at 1105 ("The challenge lies in making the social and psychological evaluations of behavior involved in legislative application of [criminal law] principles upon a practicable scale.").

38. *See* Wechsler, *supra* note 27, at 1105 ("Civilized social thought regards penal law as the ultimate weapon for diminishing the incidence of major injuries to individuals and institutions . . . . [W]hile invocation of a penal sanction necessarily depends on past behavior, the object is control of harmful conduct in the future."); *id.* at 1105–06 (remarking that "the first task of analysis [in crafting a model penal code] is to appraise and classify the major injuries with the prevention of which penal law should be concerned"); Wechsler & Michael, *supra* note 32, at 731 ("[T]he principal end of the law of homicide is the prevention of behavior which may cause death.").

39. *See id.* at 733 ("[T]he kinds of behavior to be described [by criminal law] are (1) those which it is desirable and possible to deter and (2) those which provide sufficient grounds for believing that the persons behaving in those ways may be dangerous enough in the future to warrant incapacitating or reforming them."); *accord* Wechsler, *supra* note 27, at 1105.

40. *See* Herbert Wechsler, Book Review, 49 COLUM. L. REV. 425, 426 (1949) (reviewing JEROME HALL, GENERAL PRINCIPLES OF CRIMINAL LAW (1947)) (describing Jerome Hall's "throw[ing] aside of preventive theories" of criminal law in favor of a retributive one as "a regressive major premise").

> Unless the actor is or ought to be aware of those aspects of his behavior or of the environment that give his conduct an offensive quality, the threat of sanctions cannot operate as a deterrent and the conduct does not show the individual to be a larger menace than another man [such that state-imposed incapacitation or treatment is warranted]. Criminality ought not, therefore, to depend upon external factors; it should also take account of the actor's state of mind.[41]

Wechsler's skepticism of the legal import of science, instrumentalist approach to doctrine, and belief in the core preventive charge of criminal law help illuminate the nature of the *mens rea* categories ultimately demarcated in the MPC. The Code's drafters, via published commentary, assert that "the concepts of purpose, knowledge, recklessness and negligence suffice to delineate the kinds of culpability that may be called for in the definition of specific crimes."[42] On their faces, these concepts evince bases in folk-psychological understandings of human agency and cognitive capacity.[43] Using this language, the definitions rely on differences in motive and appreciations of environments and the risks attendant with conduct in defining culpability tiers—e.g., at the top, purpose entails specific results being an individual's "conscious object," whereas knowledge involves an individual being "practically certain" that particular results will follow from his conduct, though these results are not necessarily sought; next, recklessness comprises conscious disregard for the "substantial and unjustifiable risk" that one's conduct will cause certain results, while negligence encompasses a condemnable lack of awareness of such risk.[44] The drafters, in turn, make no reference to psychological concepts or literature in their comments addressing the four mental states,[45] in contrast to comments concerning other parts of the Code.[46] The criteria appear wholly distilled (for easy consumption and implementation by legislatures)

41. Wechsler, *supra* note 27, at 1108.

42. MODEL PENAL CODE AND COMMENTARIES: OFFICIAL DRAFT AND REVISED COMMENTS 227 (1985). *See* Herbert Wechsler, *On Culpability and Crime: The Treatment of* Mens Rea *in the Model Penal Code*, 339 ANNALS AM. ACAD. POL. & SOC. SCI. 24, 27–28 (1962) ("[F]or purposes of liability only four concepts are needed to describe the kinds of culpability that may be deemed sufficient and to draw the distinctions that may usefully be drawn.").

43. MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962). For a brief description of folk psychology, see supra note 33 and accompanying text.

44. MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962).

45. MODEL PENAL CODE AND COMMENTARIES, *supra* note 42, at 229–44.

46. *See, e.g.*, *id*. at 351–64 (citing psychological literature on the cognitive effects of alcohol ingestion in a consistent manner in relation to the MPC's treatment of intoxication).

from criminal law and criminal law scholarship following an expansive survey and critical analysis,[47] with apparent primary concern for crime prevention and secondary concern for retributive principals.[48] Thus, the MPC's culpable mental states are seemingly folk-psychological products of a meticulous instrumentalist straining of criminal law without regard for scientific validity or verifiability.

B. Psychological-Behavioral Research Addressing the Model Penal Code's Mens Rea *Categories*

The MPC had immediate and lasting impact on American criminal law. Over thirty jurisdictions significantly reformed their criminal codes within thirty years of the MPC's official release and incorporated the Code's innovations to various extents.[49] With regard to culpable mental states, a 1997 survey by Professor Dannye Holley found that "twenty-two states . . . adopted the [MPC]'s scheme" while four others "emulate[ed] to a limited extent the Code's culpability concepts [but] made significant . . . departures."[50] Beyond criminal code reform, "[t]housands of court opinions have cited the [MPC] as persuasive authority for [statutory interpretation] or . . . to formulate criminal law

---

47. MODEL PENAL CODE AND COMMENTARIES, *supra* note 42, at 229–44; *see* Dubber, *supra* note 37, at 60 ("[The MPC] was as much a criminal law treatise as it was a criminal law code."); Sanford H. Kadish, *Codifiers of the Criminal Law: Wechsler's Predecessors*, 78 COLUM. L. REV. 1098, 1143 (1978) (describing the MPC as "a remarkable tour de force of analytical precision" in which "[f]our [culpable] mental states were identified and defined"); Herbert L. Packer, *The Model Penal Code and Beyond*, 63 COLUM. L. REV. 594, 594 (1963) ("[The MPC's] spirit is one of accommodation, . . . the kind of accommodation to existing institutions that results from the perception that in law one does not write on a clean slate.").

48. *See* Wechsler, *supra* note 21, at 1432, 1435 (articulating the "dominant preventive purpose of the Code," which "has important implications in determining the scope of criminality," but also noting "the demands of justice with respect to allocating blame and punishment"). According to legal scholar Sanford Kadish,

> [f]or Wechsler and his colleagues, making the law just meant at bottom that the preventive purpose of the criminal law must be sought within the constraints of the principle that punishment may not be imposed in the absence of blameworthy conduct or in disregard of the degree of blameworthiness reflected in the mental state accompanying conduct.

Kadish, *supra* note 47, at 1142. The MPC's drafters do, however, appear to address some critiques leveled against criminal law from psychology. *See* Wechsler *supra* note 27, at 1105. The Code's *mens rea* categories exclude terms psychologists considered unsound or superficial, like "malice aforethought" and "premeditation," in favor of "presumably testable phenomena such as 'conscious object' or 'knowledge,'" thereby "reducing (but not eliminating) reliance upon normative judgments." Robinson & Dubber, *supra* note 1, at 335.

49. Holley, *supra* note 25, at 229; Robinson & Dubber, *supra* note 1, at 326.

50. Holley, *supra* note 25, at 236–37. Thus, "[m]ore than half of the states have adopted mental state provisions modeled on the MPC framework," Chiesa, *supra* note 2, at 579, but "not all of the states have chosen to adhere to the Code's specific language," Holley, *supra* note 25, at 238.

doctrine."[51] In the realms of legal education and research, the MPC "provides the lingua franca of most people who teach criminal law in the United States"[52] and has "been the intellectual focus of much American criminal law scholarship since [its] promulgation."[53]

The MPC's considerable sway across criminal law, and with respect to culpable mental states in particular, makes its *mens rea* categories prime and appropriate subjects of analysis and critique.[54] In the decades since the MPC's official release, its mental state classifications have been consistently engaged, though treatments have been predominantly conceptual rather than empirical.[55] Empirical psychological-behavioral studies examining practical understanding and application of these classifications were finally conducted beginning in the early 1990s. Relevant research and findings addressing how lay people identify MPC mental states, rank them according to culpability, and employ them are briefly presented here.[56]

Three works spanning roughly thirteen years comprise what can be considered the first wave of such studies. A team led by psychologist and attorney Laurence Severance reported that subjects generally did not rank the MPC mental states in the same order of culpability as the Code or allot punishment to individuals possessing these states in legal scenarios in a manner consistent with the MPC's *mens rea* hierarchy.[57] Legal scholar Justin Levinson similarly recounted that subjects largely

---

51. Robinson & Dubber, *supra* note 1, at 327; *see* Sanford H. Kadish, *Fifty Years of Criminal Law: An Opinionated Review*, 87 CALIF. L. REV. 943, 949 (1999) ("State and federal courts commonly came to use [the MPC's] text and commentary as persuasive, if not authoritative, even in the absence of legislative reform.").

52. George P. Fletcher, *Dogmas of the Model Penal Code*, 2 BUFF. CRIM. L. REV. 3, 3 (1998)

53. Robinson & Dubber, *supra* note 1, at 320.

54. *Cf. supra* note 20 and accompanying text.

55. *See, e.g.*, Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CALIF. L. REV. 931 (2000); Kevin Cole, *Knowledge and Belief as Criminal Law Mental States*, 16 OHIO ST. J. CRIM. L. 441 (2019); Kimberly Kessler Ferzan, *Opaque Recklessness*, 91 J. CRIM. L. & CRIMINOLOGY 597 (2001); Fletcher, *supra* note 52, at 6–8 (1998); Michaels, *supra* note 20; Kenneth W. Simons, *Should the Model Penal Code's* Mens Rea *Provisions Be Amended?*, 1 OHIO ST. J. CRIM. L. 179 (2003); Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV. 463 (1992); *see also* Gainer, *supra* note 22, at 581–84 (listing several earlier conceptual critiques of the MPC *mens rea* categories).

56. This section does not critically assess the psychological-behavioral studies presented. Such an endeavor would, however, be an important avenue for future research.

57. Laurence J. Severance, Jane Goodman & Elizabeth F. Loftus, *Inferring the Criminal Mind: Toward a Bridge Between Legal Doctrine and Psychological Understanding*, 20 J. CRIM. JUST. 107, 115 (1992). These errors occurred across three different subject groups: one where simply the names of the four MPC *mentes reae* were provided, one where the names were provided and subjects had to fill in definitions themselves, and one where the names and legal definitions were provided. *Id*. at 109–10, 115.

did not assign culpability in accordance with the Code when they assessed identical conduct and harm emanating from people with different MPC mental states (purpose, knowledge, and recklessness).[58] However, legal scholar Paul Robinson and psychologist John Darley reported that subjects tended to dispense liability and punishment in line with the MPC's hierarchy of mental states when assessing legal situations differing only with regard to *mens rea*.[59]

This first wave of quantitative scholarship was uncoordinated. In contrast, a relatively concerted second wave has rolled out for approximately the past decade, primarily supported by the John D. and Catherine T. MacArthur Foundation.[60] This second wave has produced particularly provocative findings with regard to the MPC's knowing and reckless mental states, which are defined as follows:

> A person acts *knowingly* with respect to a material element of an offense when:
>
> (i) if the element involves the nature of his conduct of the attendant circumstances, he is aware that his conduct is of that nature or that such circumstances exist; and
>
> (ii) if the element involves a result of his conduct, he is aware that it is practically certain that his conduct will cause such a result . . . .
>
> A person acts *recklessly* with respect to a material element of an offense when he consciously disregards a substantial and unjustifiable risk that the material element exists or will result from his conduct. The risk must be of such a nature and degree that, considering the nature and purpose of the actor's conduct and the circumstances known to him, its disregard involves a gross deviation from the standard of conduct that a law-abiding person would observe in the actor's situation.[61]

---

58. Justin D. Levinson, *Mentally Misguided: How State of Mind Inquiries Ignore Psychological Reality and Overlook Cultural Differences*, 49 HOW. L.J. 1, 21 (2005).

59. PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY & BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW 84–96 (1995).

60. *See* Matthew R. Ginther, Francis X. Shen, Richard J. Bonnie, Morris B. Hoffman, Owen D. Jones & Kenneth W. Simons, *Decoding Guilty Minds: How Jurors Attribute Knowledge and Guilt*, 71 VAND. L. REV. 241, 241 n.2 (2018); Matthew R. Ginther, Francis X. Shen, Richard J. Bonnie, Morris B. Hoffman, Owen D. Jones, René Marois & Kenneth W. Simons, *The Language of Mens Rea*, 67 VAND. L. REV. 1327, 1327 n.τ (2014); Jones et al, *supra* note 10, at 1 n.†; Frances X. Shen, Morris B. Hoffman, Owen D. Jones, Joshua D. Greene & René Marois, *Sorting Guilty Minds*, 86 N.Y.U. L. REV. 1306, 1306 n.ε (2011); Vilares et al., *supra* note 4, at 3227.

61. MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962) (emphases added).

A team led by legal scholar Francis Shen (Shen and colleagues) reported that, for the most part, subjects "accurately sort[ed] culpable mental states into the four MPC categories" and "punish[ed] consistent with the hierarchy of severity assumed by the [Code]."[62] Yet subjects could only correctly identify conduct done knowingly and recklessly roughly 50 percent and 40 percent of the time, respectively, when deciding between the four MPC mental states and a "blameless" category.[63] And for most presentations, "there was no statistically significant difference between knowing and reckless punishment ratings. In many instances, subjects actually reversed the MPC hierarchy and punished reckless behavior more than . . . knowing ones."[64]

A team led by neuroscientist and attorney Matthew Ginther built on the work of Shen and colleagues by modifying the language used to delineate MPC *mens rea* categories for subjects.[65] This resulted in improvements in subjects' ability to accurately identify culpable mental states in legal scenarios.[66] However, "improvements [were] limited."[67] Pertinently, only 53 percent and 47 percent of subjects correctly classified situations involving knowledge and recklessness, respectively, when choosing between the four MPC *mentes reae* and a "blameless" category.[68] And when the two mental states were misidentified, they were most often confused with one another.[69] Finally, despite improvements in categorization accuracy, there were no statistically significant differences in subject punishment ratings for conduct committed knowingly or recklessly, consistent with the findings of Shen and colleagues.[70]

Another team, also led by Ginther (Ginther and colleagues (2018)), again showed subjects legal scenarios involving MPC mental states (knowledge, recklessness, and negligence), but crafted the scenarios such that there was a "best fit" *mens rea* ascription along with additional "plausible responses"—i.e., mental states that were not the

62. Shen et al., *supra* note 60, at 1354.

63. *Id*. Participants did, however, perform significantly better than chance, which would have resulted in correct responses only 20 percent of the time.

64. *Id*.

65. Ginther et al. (2014), *supra* note 60, at 1339–49.

66. *Id*. at 1352.

67. *Id*. at 1359.

68. *Id*. at 1352; *see id*. at 1330 ("[S]ubject accuracy in identifying the reckless and knowing mental states remains far below the classification accuracy for the other mental states.").

69. *Id*. at 1352.

70. *Id*. at 1360.

best fit but were plausible given the situation presented.[71] The researchers reported that subjects correctly identified conduct done knowingly, recklessly, and negligently 81 percent, 40 percent, and 53 percent of the time, respectively, when deciding among the four MPC mental states and a "blameless" option.[72] However, both reckless and negligent scenarios had "multiple plausible responses" involving different mental states, likely increasing the risk of confusion.[73] When subjects were presented with more descriptive scenario language, correct recognition of knowledge, recklessness, and negligence changed to 76 percent, 58 percent, and 63 percent, respectively, when choosing between the four MPC criteria and a "blameless" category.[74] Given these results, and the fact that many ostensibly inaccurate identifications entailed plausible responses that were not assessed as entirely incorrect, the researchers concluded that "subjects appear[ed] to have grasped the basic conceptual distinctions . . . drawn by the MPC as . . . modeled" and that "[t]he results . . . indicate that with little or no training subjects can apply the MPC framework in a manner that is largely congruent with the basic assumptions of the MPC's mental state hierarchy."[75]

Of additional note, Ginther and colleagues (2018) also reported, based on separate experiments, that subjects generally considered recklessness both necessary and sufficient for holding others criminally liable.[76] Moreover, "subjects were strongly inclined to regard recklessness as . . . sufficient . . . for liability *even when . . . instructed that 'knowledge'* [*was*] *required under the statute*."[77] These findings align with those of two other studies. Robert Beattey and Mark Fondacaro, who are both psychologists and legal scholars, recount that

71. Ginther et al. (2018), *supra* note 60 at 251–52. The authors provide examples of how they implement the best fit–plausible responses framework. One involves a fact pattern in which the uncle of a man named John "pays him one hundred dollars to bring a duffle bag across the border." The uncle is known to associate with drug dealers but claims the bag simply contains a birthday present. John's sister, who is present, says, "I don't think that's a birthday present." In response, John's uncle says, "It's just coffee," while winking towards John. "Based on the totality of the circumstances," Ginther and colleagues (2018) "think the *best fit* for these scenarios is recklessness, but knowledge and even just negligence could also be *plausible responses*, depending on a subject's judgment about the strength of the appropriate inferences." *Id*. (emphases added).

72. *Id*. at 255.

73. *Id*.

74. *Id*. at 261. A new scenario was also added in which "[a]ware[ness] of [r]isk" was signaled to entail knowledge rather than recklessness; this was correctly identified 67 percent of the time. *Id*.

75. *Id*. at 269.

76. *Id*. at 270–71.

77. *Id*. at 273.

across vignettes presented in their study, "more than 60% of participants erroneously concluded that a purposeful *mens rea* requirement had been satisfied when it was not warranted by the facts."[78] The percentage of misidentifications increased with the culpability of the mental state that was actually indicated—i.e., more subjects elevated recklessness to purpose than elevated negligence to purpose.[79] Furthermore, the percentage of such errors was correlated with the amount of harm suffered in a scenario: the more severe the harm, the higher the likelihood subjects would incorrectly ascribe the purposeful mental state requisite for criminal liability.[80] In turn, a team led by psychologist and legal scholar Pam Mueller investigating mental state attributions in civil contexts as opposed to criminal ones, reports that "not only do people regard the side effects of knowing acts as intentional when assigning liability, but they also regard the side effects of *reckless* acts as intentional when making liability judgments."[81] With regard to awareness of risk, appreciation of "even a mere 3 percent chance [of injury] was sufficient for 35 percent of participants to say that" any resulting harm was intentional.[82] Together, these studies indicate the powerful, befogging influence of moral intuitions on human assessments of the mental states of others, even when individuals are exposed to the seemingly clarifying MPC mental state delineations.[83]

The aforementioned empirical psychological-behavioral studies have expanded knowledge of real-world understandings and implementations of MPC mental state concepts.[84] They have thereby contributed to discourse on whether the MPC's *mens rea* scheme is

---

78. Robert A. Beattey & Mark R. Fondacaro, *The Misjudgment of Criminal Responsibility*, 36 BEHAV. SCI. & LAW 457, 466 (2018).

79. *Id*.

80. *Id*.

81. Pam A. Mueller, Lawrence M. Solan & John M. Darley, *When Does Knowledge Become Intent? Perceiving the Minds of Wrongdoers*, 9 J. EMPIRICAL LEGAL STUD. 859, 875 (2012).

82. *Id*. at 881.

83. *Cf*. Paul Egré, *Qualitative Judgments, Quantitative Judgments, and Norm-Sensitivity*, 33 BEHAV. & BRAIN SCI. 335, 336 (2010) ("[M]oral considerations influence . . . our qualitative evaluation of precise numerical *probabilities* and our qualitative evaluation of precise *quantities*."); Joshua Knobe, *Person as Scientist, Person as Moralist*, 33 BEHAV. & BRAIN SCI. 315, 315 (2010) ("[P]eople's judgments about whether a given action truly is morally good or bad can . . . affect their intuitions about what that action caused and what mental state the agent had.").

84. Illuminating non-empirical psychological-behavioral assessments of the MPC's *mens rea* scheme have also been undertaken over the past decade. *See, e.g.*, Kevin Jon Heller, *The Cognitive Psychology of Mens Rea*, 99 J. CRIM. L. & CRIMINOLOGY 317 (2009); James A. Macleod, *Belief States in Criminal Law*, 68 OKLA. L. REV. 497 (2016).

appropriate in the abstract and in practice. Research indicating that people have difficulty recognizing the knowing and reckless mental states and utilizing them to dispense liability and punishment as intended by the MPC is particularly poignant.[85] Such findings are alleged to undercut the normative and practical appropriateness of employing knowledge and recklessness as distinct determinants of legal culpability and even their existences as separate psychological experiences.[86] It is these issues that Vilares and colleagues endeavored to address with their cognitive neuroscience study,[87] an undertaking Jones and colleagues deem decidedly successful.[88]

II. RECENT NEUROLAW RESEARCH AND COMMENTARY

Jones and colleagues profess that a "big problem" with the MPC's knowing and reckless *mens rea* categories "is that no one knows if the legally assumed and statutorily instantiated distinction . . . reflects an actual and inherent psychological difference."[89] If such a variance exists, Jones and colleagues contend that "there would also be a difference between the brains of reckless and knowing individuals" because "anytime there is a psychological difference there must also be a brain difference."[90] Hence, if there is no neural contrast, and therefore "no detectable or meaningful psychological distinction" between the two mental states, "then widespread injustice will have followed in the wake of the [MPC], and will continue indefinitely, if unchecked."[91] That is, the legitimacy of the MPC's culpability scheme is alleged to depend on different neurological functioning underpinning the cognitive experiences of knowledge and

85. The emergent controversy surrounding knowledge and recklessness is additionally interesting because the drafters of the MPC believed that they drew "a narrow distinction between acting purposely and knowingly" relative to the "important discrimination . . . between acting either purposely or knowingly and acting recklessly." MODEL PENAL CODE AND COMMENTARIES, *supra* note 42, at 233, 236; *see* Wechsler, *supra* note 42, at 28, 29 (asserting that "[t]he discrimination between acting purposely and acting knowingly is, of course, a narrow one," and that "[t]here is a broader distinction between acting either purposely or knowingly and acting recklessly").

86. *See* Ginther et al., *supra* note 60, at 1330, 1337, 1363–64; Shen et al., *supra* note 60, at 1355–56.

87. *See* Vilares et al., *supra* note 4, at 3222 (describing the findings of Shen and colleagues and the need to further investigate the MPC's knowing and reckless mental states).

88. *See* Jones et al., *supra* note 10, at 31 (maintaining that the Vilares and colleagues study provided "empirical support" for the MPC's knowing and reckless culpable mental state classifications).

89. *Id.* at 4.

90. *Id.* at 5.

91. *Id.*

recklessness.[92] Vilares and colleagues therefore "designed and executed a brain-imaging experiment attempting to detect—for the first time—differences between mental states relevant to criminal law" (knowledge and recklessness).[93] The experiment also included a psychological-behavioral component aimed at investigating situational variables that might influence decision making.[94]

The study involved 40 participants: 20 females and 20 males with an average age of 29.[95] Each trial of the experiment involved a subject deciding whether to carry a suitcase that might contain "valuable content"—which Vilares and colleagues and Jones and colleagues refer to as "contraband" in their respective publications—through a checkpoint at which the suitcase might be searched by guards.[96] Prior to making a decision, subjects were sequentially presented with a trial's "contraband risk"—i.e., the risk that a subject was carrying contraband—and "search risk"—i.e., the risk that a subject would be searched at the checkpoint.[97] Half of the participants were randomly assigned to always see the contraband risk first while the other half always saw the search risk first.[98] Participants completed 125 trials each.[99]

Contraband risk was communicated by showing a subject "between one and five suitcases, only one of which actually contained contraband."[100] Accordingly, "the number of briefcases shown represented the risk of carrying the target suitcase with contraband."[101] If only one suitcase was presented, then the objective risk of carrying contraband was 100 percent—this was considered the "knowing situation."[102] A "reckless situation" entailed the presentation of more than one suitcase.[103] The objective contraband risk decreased as the number of suitcases shown increased.

---

92. *Id*. at 30.
93. *Id*. at 2.
94. Vilares et al., *supra* note 4, at 3223–24.
95. Jones et al., *supra* note 10, at 8 n.12.
96. *Id*. at 8; Vilares et al., *supra* note 4, at 3223.
97. Vilares et al., *supra* note 4, at 3223.
98. Jones et al., *supra* note 10, at 11. Each set of 20 participants was made up of ten males and ten females. Vilares et al., *supra* note 4, at 3223.
99. Jones et al., *supra* note 10, at 8.
100. Vilares et al., *supra* note 4, at 3223.
101. *Id*.
102. *Id*.
103. *Id*.

In turn, search risk was communicated by showing a subject "ten tunnel exits . . . , some number of which—either two, four, six, or eight—showed a guard standing prominently in the exit," indicating that the subject's suitcase would be searched if one of those tunnels was selected.[104] The objective search risk increased as the number of guards displayed increased.

Actual numerical risk profiles were never shown to participants in numeric or word form. Rather, subjects had to "deduce" the various contraband and search risks with which they were presented from the images shown.[105]

To encourage active and deliberate involvement, participants started each trial with a certain amount of virtual money ($6000) and earned more if they successfully carried a suitcase containing contraband through the checkpoint ($2000), but they lost virtual money if they were caught carrying contraband ($4000), carried a suitcase without contraband ($500), chose not to carry a suitcase ($1500), or failed to make a choice ($2500).[106] At the end of their participation, subjects received, in real money, "one percent of the payout from one trial, chosen at random," which could be "between $20.00 and $80.00."[107] Subjects never learned whether a case they carried was searched or whether such a case contained contraband; they therefore completed each trial without knowledge of the results of previous trials.[108]

The cognitive-neuroscience component of the study consisted of all subjects conducting the entirety of their involvement within an fMRI machine—"[e]ach subject was in the scanner for about 40 minutes."[109] Scans of participants' brain function were continuously obtained throughout their participation.[110] A machine learning algorithm was trained on a subset of the collected data to try and recognize and distinguish the neural functioning associated with subjects' appraisals of knowing and reckless situations.[111] Subsequently, based on this training, the algorithm was deployed to analyze additional collected data to try and predict whether a given

104. Jones et al., *supra* note 10, at 10.
105. *Id*.
106. *Id*. at 9.
107. *Id*.
108. *Id*. at 11 n.16.
109. *Id*. at 15.
110. *Id*.
111. *Id*. at 15–18.

presentation of neural functioning emanated from a subject assessing a knowing or reckless situation.[112]

The results reported by Vilares and colleagues and Jones and colleagues were produced by the algorithm processing subject brain states at the time the contraband risk was revealed—for half of the participants, this was after they had viewed the search risk, for the other half, this was before.[113] For subjects presented with the search risk first, the algorithm achieved "an average correct classification rate (CCR) of 71%"—i.e., it correctly predicted the situation (knowing or reckless) that a given subject was considering during a given trial based only on brain images 71 percent of the time.[114] For subjects presented with the contraband risk first, the algorithm achieved a CCR of 32.1 percent.[115]

The psychological-behavioral component of the study involved assessment of the impacts of risk degrees and the order of presentation of risk types (contraband and search) on decisions to carry a suitcase.[116] Broadly, "as the likelihood of a suitcase containing contraband increased, decisions to carry the suitcase decreased."[117] Similarly, "as the likelihood of being searched increased, decisions to carry the suitcase decreased."[118] Moreover, it was shown that "for identical degrees of [s]earch [r]isk . . . , seeing the search risk before contraband risk resulted in fewer decisions to carry . . . compared with seeing the [s]earch [r]isk after the [c]ontraband [r]isk."[119]

Based on the aforementioned results, Jones and colleagues offer four "immediate legal implications."[120] First and foremost, they claim that "by combining fMRI brain-imaging techniques with a machine learning algorithm [Vilares and colleagues] were able to distinguish among guilty minds."[121] With 71 percent accuracy "in some conditions," the algorithm is alleged to have correctly discerned subjects experiencing knowing mental states from those experiencing reckless mental states.[122] According to Jones and colleagues, this

112. *Id*.
113. Vilares et al., *supra* note 4, at 3224.
114. *Id*.
115. *Id*.
116. *Id*. at 3223–24.
117. *Id*. at 3223.
118. *Id*.
119. *Id*. at 3223–24.
120. Jones et al., *supra* note 10, at 21.
121. *Id*. at 20.
122. *Id*.

furnishes a "clear" affirmative answer to the question of whether the MPC's knowing and reckless *mens rea* criteria "reflect a detectable distinction between brain states."[123] In light of this, they further contend that the study "suggests that differential liability can legitimately rest . . . on there being a distinction between knowing and reckless mental states of the kind that is reflected in distinct neural activity."[124]

Second, implicating previous psychological-behavioral research alleged to indicate that subjects have difficulty identifying and discriminating knowledge and recklessness in others,[125] Jones and colleagues assert that Vilares and colleagues' results "support . . . [the] idea that jurors need more help figuring out how to distinguish knowing from reckless mental states in real cases."[126] With the study allegedly establishing knowledge and recklessness as distinct mental states worthy of differential legal treatment,[127] Jones and colleagues maintain that the just execution of the MPC's culpability scheme demands that jurors be able to accurately identify others' physically-distinct psychological dispositions so that materially dissimilar accused can be appropriately, and differentially, evaluated.[128]

Third, Jones and colleagues hold that the study "suggest[s] the [MPC] mental state categories may not be nearly as unitary as currently supposed."[129] In particular, knowledge and recklessness may be most distinct "when subjects perceive information about the presence or absence of an element of a crime after they learn information about the likelihood of being caught," and less distinct when risks are perceived in the reverse order.[130] Moreover, employing these two *mens rea* categories is alleged to perhaps be "on shakier ground" in the latter instance than the former.[131] The authors therefore ponder "whether policy-makers should consider keeping the knowing versus reckless bifurcation for some defined circumstances or types of crimes, and eliminating it for others."[132]

---

123. *Id*. at 21.
124. *Id*.
125. *See supra* notes 62–86 and accompanying text.
126. Jones et al., *supra* note 10, at 22.
127. *See supra* notes 120–124 and accompanying text.
128. Jones et al., *supra* note 10, at 22.
129. *Id*.
130. *Id*. at 23.
131. *Id*.
132. *Id*.

Finally, Jones and colleagues broadly state that the "experiment provides a concrete example of how neuroscientific methods can open up new avenues for discovering answers to some of the law's enduring questions."[133] However, due to the potentially serious experimental-design and conceptual limitations of the target study, Jones and colleagues' proffered legal implications appear deficient.

III. THE LIMITATIONS OF RECENT NEUROLAW RESEARCH AND THEIR BROADER IMPLICATIONS

This section presents notable limitations of the study by Vilares and colleagues that undermine its legal relevance. These limitations also, therefore, undermine Jones and colleagues' claims that the aforementioned study produced legally significant findings and is a model for how cognitive neuroscience can contribute to the study of law. First, interpretive limitations are presented: those that do not appear to seriously challenge Jones and colleagues' claims but still limit the explanatory power of the study by Vilares and colleagues. Second, experimental-design limitations are detailed: those indicating that the study was not designed or implemented so as to properly elicit or measure knowledge and recklessness. And third, conceptual limitations are raised: those implicating the potential for cognitive neuroscience to reveal or validate the existence of MPC mental states.

A. Interpretive Limitations

At the outset of this evaluative section, it is important to briefly note interpretive limitations of Vilares and colleagues' study that undercut claims that its reported results bear directly on knowledge and recklessness as delineated in the MPC. These constraints are not apparently fatal to Jones and colleagues' broad contention that Vilares and colleagues contributed to understanding knowledge and recklessness, but they cabin the significance of any potential contributions.

First, Vilares and colleagues' reported results stem from brain imaging data compiled at the time contraband risk was revealed to subjects.[134] But the MPC refers to individuals' psychological appreciations at the time of offensive conduct.[135] Thus, the neuroimaging data that likely most accurately correlates with knowing

133. *Id*.

134. Vilares et al., *supra* note 4, at 3224.

135. MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962).

or reckless mental states, as legally defined, are that recorded when subjects decided to carry suitcases. Jones and colleagues applying the knowledge and recklessness labels to mental states that the target study purportedly elicited therefore appears inaccurate.

Second, recklessness, as articulated in the MPC, entails "a value judgment, not . . . a fact."[136] At the moment of an act or omission, an individual disregards certain risks; these risks are only deemed "substantial" or "unjustifiable," and the disregard is only deemed "a gross deviation from the standard of conduct that a law-abiding person would observe in the actor's situation," as matters of law after offensive conduct takes place.[137] Accordingly, an individual cannot be determined to have been reckless under the MPC until after the fact by a tribunal.[138] It is therefore questionable whether Vilares and colleagues stimulated mental states that others would consider reckless. All that can be confidently claimed is that the researchers potentially induced less-than-certainty as to the presence of contraband in suitcases. This further indicates that Jones and colleagues' use of the reckless label is incorrect.

Nevertheless, Jones and colleagues' belief "that the essence of the distinction between knowing and reckless mental states . . . reflects different probabilities" appears apt.[139] As a result, discerning whether different neural functioning attends psychological appreciations of certainty and less-than-certainty seems relevant to determining whether such variance corresponds with mental states deemed knowing and reckless. Given the importance of probability assessment to knowledge and recklessness, this Essay entertains Jones and colleagues' inaccurate designations.

The aforementioned limitations restrain the extent to which Vilares and colleagues could have contributed to understanding the MPC's knowing and reckless culpability criteria, but do not apparently

136. Wechsler, *supra* note 42, at 30.

137. MODEL PENAL CODE § 2.02(2)(c) (AM. LAW INST., Proposed Official Draft 1962). To be sure, knowledge can entail a value judgment as well with its call to determine whether it was "*practically certain* that [an individual's] conduct [would] cause . . . a result." *Id.* at § 2.02(2)(b) (emphasis added). But Vilares and colleagues appear to obviate this issue by only referring to mental states potentially entailing certainty as knowledge. Vilares et al., *supra* note 4, at 3223.

138. *See* Packer, *supra* note 47, at 601–02 (noting that recklessness "require[s] a standard to be announced for [its] application to specific cases" and that, after conduct involving disregard for risk, "it remains to articulate a standard for judging when a risk is 'substantial and unjustifiable'").

139. Jones et al., *supra* note 10, at 29–30.

prevent them from doing so. The limitations forthcoming, however, are perhaps more impactful in this regard.

B. Experimental-Design Limitations

Cognitive neuroscience studies are only as good as their predicate psychological-behavioral components.[140] Target mental states and behaviors must be properly evoked—in terms of both presence and discreteness—in order for corresponding neural functioning to be reliably discerned. That is, "[a] minimal requirement for a well-designed [cognitive neuroscience] study is that subjects actually perform [what] the experimenter believes they are performing,"[141] and the "ability to make judgments about brain maps and their relationship with psychological processes is limited by how strictly a task . . . tests the mental process of interest."[142] The tasks employed by Vilares and colleagues do not appear properly designed to elicit and isolate the knowing and reckless mental states or allow for the identification of the states' neurological correlates, nor do they seem to accomplish these aims.[143]

It is unclear whether or to what extent participants in the target study experienced knowing or reckless mental states. Subjects were presented with what the researchers termed "knowing *situations*" and "reckless *situations*" in which the *objective* chances of carrying contraband were 100 percent and less-than-100 percent, respectively.[144] But objective risk profiles were never displayed to participants in numerical or word form and participants were never required to record their understandings of contraband risk. This implicates the study's

---

140. Pertinently, legal scholar Stephen Morse argues that

> behavioral neuroscience is largely dependent on psychology. Neuroscientists do not go on expensive fishing expeditions without knowing what they are hoping to catch. Instead, they have already identified some psychological or behavioral trait or condition . . . that interests them . . . . Such identification depends on those behavioral conditions already being well-characterized and operationalized.

Stephen J. Morse, *Is Executive Function the Universal Acid?*, 15 CRIM. L. & PHIL. 1, 3–4 (forthcoming 2021); *see* Roskies, *supra* note 8, at 71 ("Activation studies with fMRI are truly a marriage between neuroscience and psychology, drawing as much upon an understanding of cognition and cognitive approaches as upon knowledge of the brain.").

141. *Id.* at 56.

142. Emily Bell & Eric Racine, *Enthusiasm for Functional Magnetic Resonance Imaging (fMRI) Often Overlooks Its Dependence on Task Selection and Performance*, 9 AJOB NEUROSCIENCE 23, 24 (2009).

143. *Cf.* Roskies, *supra* note 8, at 53 ("Selection of the tasks involved in a study is critically important, strongly affecting its interpretability and outcome.").

144. Vilares et al., *supra* note 4, at 3223 (emphases added).

construct validity—i.e., the degree to which it actually measures what it aims to measure.[145]

Validly decoding the neural correlates of psychological phenomena via neuroimaging experiments requires that "brain activity is recorded in circumstances in which there is an independent way to determine what psychological state the subject is in."[146] Jones and colleagues presume that those assessing knowing situations experienced knowledge with regard to the presence of contraband based on the number of suitcases presented and behavior allegedly consistent with such knowledge.[147] Previous psychological-behavioral research, however, indicates that people have difficulty correctly identifying and distinguishing knowledge and recklessness in others and knowing and reckless situations: experiment subjects consistently misclassified individuals experiencing knowledge and recklessness and situations involving knowledge and recklessness; these misclassifications often involved subjects confusing knowledge with recklessness and vice versa.[148] In addition, subject behavior in the Vilares and colleagues study when presented with knowing situations was consistent with appreciating high contraband risk, not necessarily a 100-percent chance of carrying contraband specifically. Finally, while the computer algorithm semi-reliably discerned neural functioning corresponding with appraising knowing situations from that corresponding with appraising reckless situations generally, in certain circumstances, it did not dependably discern neurological activity associated with assessing knowing situations from that associated with assessing reckless situations involving two suitcases—i.e., situations entailing an objective contraband risk of 50 percent. The computer algorithm achieved an average correct classification rate of only 55

145. *See* MADHU VISWANATHAN, MEASUREMENT ERROR AND RESEARCH DESIGN 68 (2005) ("Construct validity is an umbrella term that asks the basic question, Does a measure measure the construct it aims to measure."); Drew Westen & Robert Rosenthal, *Quantifying Construct Validity: Two Simple Measures*, 84 J. PERSONALITY & SOC. PSYCH. 608, 608 (2003) ("Construct validity is . . . at the heart of any study [that] use[s] a measure as an index of a variable that is not itself directly observable . . . . [Without] construct validity, results obtained . . . will be difficult to interpret.").

146. Gideon Yaffe, *Mind-Reading by Brain-Reading and Criminal Responsibility*, *in* PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE 144 (Dennis Patterson & Michael S. Pardo eds., 2016).

147. Jones et al., *supra* note 10, at 12. Jones and colleagues note that they "cannot claim that all subjects actually knew what they clearly should have known," but nevertheless presume such knowledge throughout their essay. *Id.* at 30.

148. *See supra* notes 62–86 and accompanying text.

percent in this regard.[149] This near indistinguishability reveals that subjects consistently inaccurately evaluated knowing situations and two-suitcase reckless situations. Accordingly, brain images taken from subjects contemplating knowing and reckless situations do not seem to represent the neurological underpinnings of knowledge and recklessness.

Even if Vilares and colleagues reliably elicited knowledge and recklessness with regard to contraband, they do not appear to have done so discretely enough to discern corresponding neural activity.[150] Vilares and colleagues implemented a decision-making task: participants were charged with deciding whether to carry a suitcase through a checkpoint. While the exercise was divided into three events (assessing contraband risk, assessing search risk, and making a decision), at no point was it incumbent on subjects to appraise individual risk factors outside the confines of the global decision-making endeavor. Pertinently, subjects were never compelled to formulate distinct knowing or reckless mental states with specific regard to the presence of contraband. Rather, they needed to quickly assimilate and integrate sequentially presented variables in light of a previously-declared incentive structure to decide whether to carry a suitcase. The overarching decision-making task involved a litany of continuous and intermittent cognitive sub-tasks, which conceivably polluted one another and prevented discernment of the distinct neural functioning of any particular mental state or appreciation. Resultantly, the only cognitive neuroscience findings possible by Vilares and colleagues appear to relate to neurological activity supporting the decision-making task generally or buttressing the task at certain points.[151]

This deduction is supported by the fact that the computer algorithm in question was only able to reliably distinguish the neural functioning of participants appraising knowing situations in one circumstance: when contraband risk was displayed after search risk. It

149. Vilares et al., *supra* note 4, at 3224.

150. Typically, "[i]n cognitive neuroscience research, . . . [a]n experimental task is designed to attempt *isolation* of the relevant psychological process that is being assessed." Brown & Murphy, *supra* note 8, at 1142 (emphasis added). In turn, "[t]he resulting captured data and constructed images are interpreted as the neural correlates of that [psychological process]." *Id.* at 1143.

151. *Cf.* Bell & Racine, *supra* note 142, at 23 ("Since fMRI operates under the premise that measured brain activity is a reflection of areas of the brain engaged in or involved in carrying out a task, conventional fMRI and the resulting brain images are bound to the context of the task."); SATEL & LILIENFELD, *supra* note 7, at 22 ("When it comes to interpreting the results of imaging studies, context is everything.").

seems unlikely that viewing contraband risk and search risk in alternate orders made subjects apprehend the same objective contraband risk meaningfully differently in terms of harboring knowing or reckless mental states. It is perhaps more plausible that differing sequences uniquely influenced overarching decision-making. As evidenced by the psychological-behavioral data, subjects broadly became less willing to carry a suitcase when search risk was presented first.[152] This condition might have enhanced sensitivity to search risk and loss, altered the weight ascribed to contraband risk, or uniquely influenced the formation of a decision through some other effect or combination of effects that produced novel, observable neurological activity. And any impacts may have been most potent in situations involving high objective contraband risks since the computer algorithm was able to semi-reliably discern neural functioning associated with subjects assessing knowing situations. Moreover, as noted above, the computer algorithm could not dependably distinguish brain activity correlated with the appraisal of knowing situations from that correlated with the appraisal of two-suitcase reckless situations.[153] This additionally supports the inference that the program did not discern a neuromarker of knowing mental states, but instead discerned conditional neural functioning associated with decision-making cognition induced by the presentation of high objective contraband risk following the presentation of search risk.

For the aforementioned reasons, Vilares and colleagues' study does not appear properly designed to evoke knowledge or recklessness as to the presence of contraband or to identify specific neural functioning correlated with these mental states. And the experiment does not appear to have been successful in either of these pursuits. Consequently, it seems inaccurate for Jones and colleagues to label neural functioning correlated with appraising knowing and reckless situations as representing the physical underpinnings of knowledge and recklessness. It also seems inaccurate to claim that Vilares and colleagues distinguished knowing mental states from reckless mental states in neurological terms or supported the elicitation of knowledge being substantially impacted by the order in which contraband and search risk were presented.

152. Vilares et al., *supra* note 4, at 3223–24.

153. *Id.*

C. Conceptual Limitations

Beyond apparent issues with stimulating and measuring desired psychological and neurological outputs, Vilares and colleagues—who predominantly employ cognitive neuroscience methods—do not seem to have suitably conceived their study to address the existence or propriety of the MPC's knowing and reckless culpability criteria. This is due to the language of the Code, the nature of mental states, and the current meager psychological-behavioral understanding of knowledge and recklessness.

The MPC's *mens rea* categories are defined in psychological-behavioral terms: mental states are the only stipulations.[154] And the reality of mental states is fundamentally a psychological-behavioral matter. Only psychological-behavioral manifestations offer base evidence of a mental state's existence; neural functioning is trivial in this regard if untethered to adequately established psychological-behavioral reference points.[155] Accordingly, a mental state's existence must be established at the psychological-behavioral level before any legitimate investigation of its physical substrates can be carried out—i.e., before cognitive neuroscience investigations can provide interpretable data.[156] If the existence of a mental state is ascertained, only then can cognitive neuroscience be employed to detect corresponding neurological activity—non-existent mental states have no physical underpinnings.[157] And only when corresponding brain activity is discerned to a reasonable degree—such that specific neural

---

154. MODEL PENAL CODE § 2.02(2) (AM. LAW INST., Proposed Official Draft 1962).

155. As put by philosophers Tuomas Pernu and Nadine Elzein,

> [i]n assessing . . . neuroscientific data, we are engaged in the project of connecting such data to psychological and behavioral data. Although we are easily led into thinking that the former is somehow the more primitive and fundamental of the two[,]. . . it is in fact on the basis of the psychological and behavioral data that we draw conclusions about the function of the neural features that are being studied.

Tuomas K. Pernu & Nadine Elzein, *From Neuroscience to Law: Bridging the Gap*, 11 FRONTIERS PSYCH. 1, 11 (2020).

156. *See* Christopher R. Fetsch, *The Importance of Task Design and Behavioral Control for Understanding the Neural Basis of Cognitive Functions*, 37 CURRENT OP. NEUROBIOLOGY 16, 16 (2016) ("If we want to understand how neural activity gives rise to our sensations, thoughts, memories, and decisions, we must come to grips with the fact that everything we know about such internal processes can only be inferred by observing external behavior."); *see also* Yael Niv, *On the Primacy of Behavioral Research for Understanding the Brain*, *in* CURRENT CONTROVERSIES IN PHILOSOPHY OF COGNITIVE SCIENCE 139 (Adam J. Lerner, Simon Cullen & Sarah-Jane Leslie eds., 2020) (arguing "that pure behavioral research is not only critical for understanding the mind—it is also the cornerstone of understanding how cognitive processes are implemented in the brain").

157. *See* Pernu & Elzein, *supra* note 155, at 12 ("To establish a reliable connection between behavioral and neural data, we need to rely on clearly defined behavioral variables.").

functioning might be accepted as a reliable indicator of a particular psychological experience—can neurological markers potentially be employed in place of psychological-behavioral indicators to determine whether individuals experience a given mental state in certain circumstances. Thus, in order for cognitive neuroscience to play a productive part in assessing mental states, including those of the MPC, their actuality must have already been established by psychological-behavioral research.[158]

Psychological-behavioral studies can also more directly contribute to discussions concerning the practical, moral, and legal appropriateness of the MPC's culpability classifications. By evaluating psychological appreciations of circumstances and behavior, researchers can determine whether *mentes reae* are stimulated, recognized in others, or incorporated into accountability assessments as the MPC presumes. Findings that contradict the MPC's presumptions might influence appreciations of the Code and engender change. Identifying neural functioning correlated with psychological conditions, on the other hand, contributes nothing on its own to practical or moral understandings of these conditions or whether they should be employed in criminal law doctrine. Only once cognitive neuroscience sufficiently links distinct neurological activity to particular mental states can brain imaging studies potentially be employed to investigate when and how such mental states are elicited and thereby possibly produce data of practical, moral, or legal importance.

The preceding exposition shows that even if it is the case, as Jones and colleagues assert, that "anytime there is a psychological difference there must also be a brain difference,"[159] it does not follow that "the moral legitimacy of the [MPC]'s taxonomy of culpable mental states . . . depends on whether those mental states actually correspond to different brain states in the way the [MPC] categorization assumes."[160] If psychological difference is concomitant with neurological difference, then the MPC's *mens rea* taxonomy depends just as much on the former as the latter. Accepting Jones and colleagues' contention means that neurological difference could be

158. *See supra* note 155 and accompanying text.

159. Jones et al., *supra* note 10, at 5.

160. *Id*. at 30; *cf*. Pernu & Elzein, *supra* note 155, at 11 ("Given that . . . neuropsychological reasoning, imaging studies in particular, [generally] proceeds" from psychology-behavior to neural functioning, with the former determining the salience of the latter, "it is very problematic to start basing our moral and legal judgments on neuroscientific data").

inferred from psychological difference. And given that psychological-behavioral manifestations, not brain activity, are the sole means of determining whether mental states exist and are the lynchpins of any potential contributions from cognitive neuroscience to understanding psychological phenomena, Jones and colleagues appear to inappropriately extol neurocentrism when the primacy of psychological-behavioral research seems apt.[161] Brain imaging data in accordance with psychological-behavioral data is largely neuroredundant in terms of the reality or presence of mental states,[162] while brain imaging data in conflict with psychological-behavioral data should be considered suspect.

For example, suppose psychological-behavioral studies indicate that individuals experience knowledge as articulated by the MPC. If Vilares and colleagues are unable to distinguish unique neural functioning associated with knowledge, this does not challenge the existence of knowledge, which can only be initially established by psychological-behavioral manifestations, not neurological activity. Rather, this implicates the neuroscience technology and methods employed by Vilares and colleagues or understandings of the relationship between neural functioning and mental states. In turn, if knowledge is belied by psychological-behavioral research but Vilares and colleagues distinguish unique neural functioning associated with subjects appraising knowing situations, this does not represent conflicting findings as to the existence of knowledge. Instead, the cognitive neuroscience results should stoke interest in what mental activity is correlated with the novel brain functioning recorded. Moreover, in both of the aforementioned situations, Vilares and colleagues' results have no practical, moral, or legal salience because they merely revealed neural functioning associated with target mental states and situations—the findings simply provide a translation of psychology-behavior to neuroscience.[163]

---

161. SATEL & LILIENFELD, *supra* note 7, at xix ("'[N]eurocentrism' [is] the view that human experience and behavior can be best explained from the predominant or even exclusive perspective of the brain. From this popular vantage point, the study of the brain is somehow more 'scientific' than the study of human motives, thoughts, feelings, and actions.").

162. "'[N]euroredundancy' denote[s] things we already knew without brain scanning." *Id.* at 22.

163. *Cf.* RUSSELL A. POLDRACK, THE NEW MIND READERS: WHAT NEUROIMAGING CAN AND CANNOT REVEAL ABOUT OUR THOUGHTS 68 (2018) ("One way to think about brain decoding [i.e., identifying the neural correlates of cognition and behavior] is that it is trying to translate between two languages: The natural language of humans, and the biological 'language' of thought in the brain.").

As it stands, psychological-behavioral studies indicate that people have trouble correctly identifying and distinguishing knowledge and recklessness and do not reliably accord punishment in line with the MPC's culpability scheme.[164] This potentially calls into question whether individuals experience knowledge and recklessness as the Code supposes and, even if so, whether the two mental states are proper culpability markers. However, it does not appear as though any empirical psychological-behavioral studies have addressed the base existence of knowledge or recklessness,[165] and while quantitative research, like fMRI studies, can help inform determinations of the propriety of MPC mental states, it cannot resolve such normative issues.[166]

Vilares and colleagues do not appear to address either of the aforementioned matters (the reality of knowledge and recklessness or their aptness as signals of criminality). The experiment did not reveal the existence of knowledge or recklessness because it did not reliably induce or ascertain subject mental states.[167] By investigating physical manifestations without determining mental reality, Vilares and colleagues put the cart before the horse and seemingly descried the neural functioning of psychological enigmas. The study also contributes little to assessing the practical, moral, or legal appropriateness of knowledge and recklessness as indicia of blameworthiness because it is silent as to their actuality and offers no insight into how subjects experience or perceive these mental states or utilize them to dispense accountability.

In light of the foregoing analysis, Jones and colleagues' claim that the cognitive neuroscience methods utilized by Vilares and colleagues empirically identified knowledge and recklessness as actual, distinct mental states appears dubious. Jones and colleagues' additional contentions that Vilares and colleagues' findings support maintaining the MPC's culpability scheme and putting additional focus on ensuring

164. *See supra* notes 62–86 and accompanying text.

165. This suggests a fruitful avenue for future research.

166. *See, e.g.*, Octavio S. Choi, *What Neuroscience Can and Cannot Answer*, 45 J. AM. ACAD. PSYCHIATRY & L. 278, 284 (2017) ("[A]lthough neuroscience can inform, it will never be able to answer ultimate legal questions of culpability and desert. Such determinations are essentially moral judgments that require understanding behaviors and mental states against the backdrop of cultural norms."); Grigori Guitchounts, *The Existential Crisis in Neuroscience*, NAUTILUS (Jan. 23, 2020), http://nautil.us/issue/81/maps/an-existential-crisis-in-neuroscience [https://perma.cc/5RRV-LXVY] ("[S]cience gathers facts about the world, but it is humans who describe it and try to understand it. And these processes require filtering the raw data through a personal sieve, sculpted by the language and culture of our times.").

167. *See supra* notes 140–153 and accompanying text.

jurors understand and implement this scheme correctly seem similarly questionable.

CONCLUSION

The MPC's culpability scheme, encompassing four mental states that must accompany offensive conduct in order for criminal liability to attach, is the most influential and pervasive framework for assigning criminal responsibility.[168] Consequently, the promulgation of equitable criminal law requires investigating the reality of the *mentes reae* the Code delineates and how they are actually experienced, perceived, and utilized to assess blameworthiness and accord punishment. Unfortunately, contrary to the assertions of Jones and colleagues, Vilares and colleagues appear to provide little insight in these regards.

Vilares and colleagues aimed to identify and distinguish neural functioning correlated with the MPC's knowing and reckless mental states in an effort to help determine whether these states actually exist and their appropriateness as discrete culpability markers.[169] Jones and colleagues, in turn, deem the aforementioned enterprise successful and of significant legal importance.[170] However, Vilares and colleagues do not seem to have adequately designed their study tasks to reliably stimulate knowledge and recklessness in subjects or, even if they did, the study does not discern brain activity uniquely associated with these mental states.[171] Indeed, the researchers appear to have failed to accomplish either of these aims.[172] More broadly, cognitive neuroscience is ill-suited to aid in ascertaining the existence of mental states or, to a certain extent, producing findings practically, morally, or legally relevant to whether specific mental states should be determinants of criminal responsibility.[173] Psychological-behavioral methods, on the other hand, are essential in these regards.[174] Accordingly, Vilares and colleagues trivially contributed to understandings of knowledge and recklessness and the propriety of the MPC's culpability criteria, if they contributed at all.[175] Rather than "provid[ing] a concrete example of how neuroscientific methods can

168. *See supra* notes 2, 20 and accompanying text.
169. Vilares et al., *supra* note 4, at 3222–23.
170. *See supra* notes 120–133 and accompanying text.
171. *See supra* notes 140–153 and accompanying text.
172. *Id.*
173. *See supra* notes 154–167 and accompanying text.
174. *Id.*
175. *Id.*

contribute information relevant to legal policy," as Jones and colleagues advocate,[176] Vilares and colleagues should perhaps serve as a cautionary illustration of deficient experimental design and the limited contributions that cognitive neuroscience can make to understanding and assessing the MPC's culpable mental states.[177] In the future, the following should be dutifully appreciated when cognitive science is brought to bear on the Code's psychological delineations: psychological-behavioral research is primary, and it must be adequately carried out before neuroscience can provide data of any legal import.

176. Jones et al., *supra* note 10, at 31.

177. According to Satel and Lilienfeld,

> [t]he [primary] job of [cognitive] neuroscience is to elucidate the brain mechanisms associated with mental phenomena, and when technical prowess is applied to the questions it can usefully address, the prospects for conceptual breakthroughs and clinical advances are bountiful. Asking the wrong questions of the brain . . . is at best a dead end and at worst a misappropriation of the mantle of science.

SATEL & LILIENFELD, *supra* note 7, at 152.