

Internal and External Challenges to Culpability

Stephen J. Morse*

“[E]ven a dog distinguishes between being stumbled over and being kicked.”¹

The thesis of this Article is simple: As long as we maintain the current folk psychological conception of ourselves as intentional and potentially rational creatures, as people and not simply as machines, mental states will inevitably remain central to ascriptions of culpability and responsibility more generally. It is also desirable. Nonetheless, we are in a condition of unprecedented internal challenges to the importance of mental states in the context of mental abnormalities and of external challenges to personhood and agency based on the new behavioral neuroscience and genetics. The latter challengers argue that the central role the criminal law gives to mental states is deeply misguided.

I begin with the law’s conception of the person as a folk psychological agent who can potentially be guided by reason. Then I canvas the internal challenges to the importance of mental states, rooted in a trilogy of United States Supreme Court opinions: *Montana v. Egelhoff*,² *Clark v. Arizona*,³ and

* Ferdinand Wakeman Hubbell Professor of Law, Professor of Psychology and Law in Psychiatry, and Associate Director, Center for Neuroscience and Society, University of Pennsylvania. Judge Morris Hoffman co-authored the original mens rea material from which the present version is derived, *The Uneasy Entente Between Legal Insanity and Mens Rea: Beyond Clark v. Arizona*, 97 J. CRIM. L. & CRIMINOLOGY 1071 (2007), and Professor Richard Bonnie co-authored the original insanity defense material from which the present version is derived, *Abolition of the Insanity Defense Violates Due Process*, 41 J. AM. ACAD. PSYCHIATRY L. 488 (2013). Ed Greenlee and Andy Lang furnished fabulous research assistance. This article is also adapted in part from my prior works: *The Neuroscientific Non-Challenge to Meaning, Morals, and Purpose*, in NEUROEXISTENTIALISM: MEANING, MORALS, AND PURPOSE IN THE AGE OF NEUROSCIENCE 333 (Gregg D. Caruso & Owen J. Flanagan eds., 2018); *The Inevitable Mind in the Age of Neuroscience*, in PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE 29 (Dennis Patterson & Michael S. Pardo eds., 2016); *Neuroscience, Free Will, and Criminal Responsibility*, in FREE WILL AND THE BRAIN: NEUROSCIENTIFIC, PHILOSOPHICAL, AND LEGAL PERSPECTIVES 251 (Walter Glannon ed., 2015); *Criminal Responsibility and the Disappearing Person*, 28 CARDOZO L. REV. 2545 (2007); and *Fear of Danger, Flight from Culpability*, 4 PSYCH. PUB. POL’Y & L. 250 (1998).

1. OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 3 (Dover Publ’ns 1991) (1881).
2. 518 U.S. 37 (1996).
3. 548 U.S. 735 (2006).

Kahler v. Kansas.⁴ In each, the Court permitted limitations on the extent to which defendants could introduce evidence of mental abnormalities to avoid convictions. I conclude that all these decisions are misguided and that such limitations should not be adopted legislatively or judicially and should be rolled back by legislation whenever possible. My goal is not to criticize the formal legal reasoning of the opinions, although I do so in passing, but to engage at the level of policy. The action is now in the courts and legislatures.

Next, I address the newer, broader challenges to personhood, agency, and responsibility that are fueled by alleged advances in behavioral neuroscience and genetics. Some of these are quite radical. They may even turn out to be correct, but at present, there is no conceptual or empirical reason to believe that they are true. Moreover, there is certainly insufficient reason to jettison notions of criminal responsibility that have been developing for centuries and to adopt instead the proposed, radical conception of justice. I do not argue for any particular categorization or hierarchy of mens rea terms,⁵ or for any particular form of an affirmative defense of legal insanity. Nevertheless, culpability, as expressed in mental state requirements (including action), is central to our value as moral agents.

I. THE CHALLENGE TO PERSONHOOD, ACTION, AND RESPONSIBILITY

As I type the words of this chapter, I have an experience that virtually every neurologically intact human being takes for granted: the subjective experience of first-person agency, the experience of mental causation, that my bodily movements and thoughts are caused, roughly speaking, by my intentions.⁶ To the best of our knowledge, only human beings potentially have a fully developed capacity to act for reasons. This description sounds like Cartesian dualism—the notion that we have an immaterial mind or soul that is somehow in causal relation with our physical body and that causes it to move as the mind directs. But I fully accept that we inhabit a thoroughly

4. 140 S. Ct. 1021 (2020).

5. See Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV. 463 (1992) (providing a critique of the present hierarchy and a new proposal); see also Kimberly Kessler Ferzan, *Don't Abandon the Model Penal Code Yet! Thinking Through Simons's Rethinking*, 6 BUFF. CRIM. L. REV. 185 (2002). See generally Paul H. Robinson, *A Brief History of Distinctions in Criminal Culpability*, 31 HASTINGS L.J. 815 (1980).

6. I am not suggesting that all bodily movements and thoughts are so caused. Many bodily movements are simply mechanistically caused, such as reflexes, and many thoughts simply spring to mind without any conscious intention to produce them. Some behavior, such as habitual gestures or verbal “tics,” does not seem intentional, but neither is it purely mechanistically produced. One can intentionally bring such movements under conscious intentional control. Such behavior is reason responsive.

material physical universe in which all phenomena are consistent with physical laws and causal closure. The latter simply means that every phenomenon in the universe is or is derived from physical matter and there are no mysterious entities. In particular, human action and consciousness are produced by the brain, a material organ that works according to biophysical laws. At present, however, we do not have a clue about how the brain enables the mind, or about how action and consciousness are possible.⁷ Understanding how the brain enables the mind would revolutionize our understanding of biological processes and the nature of personhood,⁸ but such understanding may not be possible.⁹

Although action and consciousness are scientific and conceptual mysteries,¹⁰ they are at the heart of both commonsense, “folk psychology,” and the conception of the person inherent in judgments about responsibility and culpability. Folk psychology is a generic term for all psychological theories that in part explain human action by mental states such as desires, beliefs, and intentions.¹¹ No folk psychological theory thinks human action is fully causally explained by mental states, but all agree that mental states are central to full causal explanation of human action. The capacity for intentional movement and thoughts—the capacity for agency—is a central aspect of personhood and is integral to what it means to be a responsible person. We act because we intend. Responsibility judgments depend on the mental states that produce and accompany our bodily movements. This is how we think about ourselves, and this is the concept of the person that morality and law both reflect. Law and morality as action-guiding normative systems of rules are useless, and perhaps incoherent, unless one accepts this view of personhood. This explains why the law is and must be a thoroughly folk-psychological institution.¹²

Virtually everything for which we deserve to be praised or blamed and rewarded or punished is the product of mental causation and, in principle,

7. Ralph Adolphs, *The Unsolved Problems of Neuroscience*, 19 TRENDS IN COGNITIVE SCIS. 173, 174 (2015); MATTHEW COBB, THE IDEA OF THE BRAIN: THE PAST AND FUTURE OF NEUROSCIENCE 4 (2020); PAUL R. MCHUGH & PHILLIP R. SLAVNEY, THE PERSPECTIVES OF PSYCHIATRY 11–12 (2d ed. 1998).

8. MCHUGH & SLAVNEY, *supra* note 7, at 12.

9. See generally COLIN MCGINN, THE MYSTERIOUS FLAME: CONSCIOUS MINDS IN A MATERIAL WORLD (1999) (arguing that understanding consciousness is impossible for creatures with our limited intellectual capacities).

10. See generally ROBERT AUDI, ACTION, INTENTION, AND REASON 1–4 (1993) (describing the basic philosophical divisions in each of the four major problem areas in action theory).

11. This general meaning must be distinguished from bits of common wisdom that may or may not be true.

12. See Katrina L. Sifferd, *In Defense of the Use of Commonsense Psychology in the Criminal Law*, 25 LAW & PHIL. 571 (2006).

responsive to reason. Machines may cause harm, but they cannot do wrong, and they cannot violate expectations about how we ought to live together. Only people can violate expectations of what they owe each other, and only people can do wrong. Machines do not deserve praise, blame, reward, or punishment. Machines do not deserve concern and respect simply because they exist. They do not have a sense of past, present, and future. These concepts apply only to potentially acting, intentional agents.

Suppose, however, that our conscious or potentially conscious intentions are not genuinely causal or seldom are so. To use the title of a book by an eminent psychologist, suppose that our “conscious will” is just an illusion.¹³ Ordinary notions of action and agency are allegedly under attack from psychology and neuroscience,¹⁴ a critique that some legal scholarship has begun to embrace.¹⁵ If this is correct, the potential normative implications are profound. Most centrally, if conscious will is an illusion, then concepts of responsibility and desert may be equally illusory or at least inapplicable in most cases of human activity. Perhaps no one really deserves anything, and human beings are morally indistinguishable from machines. Although many people think that the implications of a thoroughly physical worldview are profound,¹⁶ I shall argue below that one can fully and consistently accept a material, matter-first worldview and also accept traditional notions of personhood, action, responsibility, and desert. But first let us consider the internal challenges to the importance of culpability in the context of mental abnormalities.

13. DANIEL M. WEGNER, *THE ILLUSION OF CONSCIOUS WILL* (2002).

14. Denise Park introduces a group of papers concerned with scientific study of will and writes that the premise of all of them is:

There are mental activations of which we are unaware and environmental cues to which we are not consciously attending that have a profound effect on our behavior and that help explain the complex puzzle of human motivations and actions that are seemingly inexplicable, even to the individual performing the actions.

Denise C. Park, *Acts of Will?*, 54 AM. PSYCH. 461, 461 (1999).

15. See, e.g., Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts*, 87 MINN. L. REV. 269 (2002).

16. See JANET RADCLIFFE RICHARDS, *HUMAN NATURE AFTER DARWIN: A PHILOSOPHICAL INTRODUCTION* 15–23 (2000). Richards argues that the thoroughly material view of people exemplified by Darwin’s theory appears to challenge traditional notions of what is most distinctive about humans. *Id.*

II. THE INTERNAL CHALLENGES: MENTAL ABNORMALITY, MENS REA, AND CULPABILITY

This section discusses three doctrinal contexts in which the emphasis on mens rea and responsibility generally are challenged: the use of voluntary intoxication to negate mens rea, the use of mental disorder to negate mens rea, and the existence of an independent, affirmative defense of legal insanity. It uses the Supreme Court jurisprudence in these areas to frame the issue.

A. Voluntary Intoxication

Montana v. Egelhoff is the leading Supreme Court precedent.¹⁷ I start with the factual background because the drama of the facts often drives policy. In July 1992, while camping in the woods of Montana, James Allen Egelhoff met and became friends with Roberta Pavola and John Christenson, who were also camping.¹⁸ On July 12, they spent the day and evening drinking heavily together.¹⁹ Around midnight, police officers discovered Christenson's station wagon in a ditch along a highway.²⁰ Pavola and Christenson were dead in the front seat, each with a single gunshot in the head.²¹ Egelhoff was lying in the rear of the car, yelling obscenities.²² The officers found Egelhoff's .28 caliber handgun on the floor near the brake pedal.²³ There were four loaded rounds and two empty casings.²⁴ Egelhoff had gunshot residue on his hands.²⁵ More than an hour after Egelhoff was taken into custody, his blood alcohol content was .36.²⁶

Egelhoff was charged with deliberate homicide, which Montana defines as "purposely or knowingly" causing the death of another human being.²⁷ "Purposely" and "knowingly" are both subjective. Egelhoff claimed that as the result of extreme intoxication, admittedly voluntarily induced, he lacked the physical ability to commit homicide.²⁸ Egelhoff was permitted to introduce evidence of his intoxication at trial.²⁹ But pursuant to a Montana

17. 518 U.S. 37 (1996).

18. *Id.* at 40.

19. *Id.*

20. *Id.*

21. *Id.*

22. *Id.*

23. *Id.*

24. *Id.*

25. *Id.*

26. *Id.*

27. MONT. CODE ANN. § 45-5-102 (2021).

28. *Egelhoff*, 518 U.S. at 41.

29. *Id.*

statute prohibiting a defendant from using intoxication evidence to negate mental states,³⁰ the trial court instructed the jury that it could *not* consider Egelhoff's intoxication in determining whether he purposely or knowingly killed the victims, as required by the definition of the crime.³¹ Egelhoff was convicted, and he appealed on the ground that the intoxication statute violated due process because it prevented the jury from considering relevant evidence and thus relieved the State from proving all the elements of the crime beyond a reasonable doubt.³² The Montana Supreme Court agreed that the statute violated due process and reversed.³³ Montana appealed to the U.S. Supreme Court.³⁴

If James Allen Egelhoff killed when he was genuinely in a state of alcohol-induced "unconsciousness," he did not kill purposely or knowingly as Montana law defines these mens reas. Perhaps he killed recklessly if he was consciously aware before or during his drinking binge that there was a substantial and unjustifiable risk that he would become homicidal if he drank to the stage of unconsciousness. Or perhaps he was not aware that his drinking created this risk, but he should have known that this risk existed. If so, he is guilty of killing negligently. But if Egelhoff did not kill purposely or knowingly—a factual issue usually left fully to the jury—then according to standard principles of culpability and desert, he does not deserve to be punished for killing with one of these two heinous mental states. He nevertheless demonstrated that he is capable of multiple homicides when drunk. Egelhoff is undoubtedly a dangerous agent.

Montana's statute expresses moral condemnation of behaving badly when drunk. Aristotle, for example, thought that a person who did harm when drunk was undoubtedly culpable.³⁵ But getting drunk is one wrong, and whatever else an agent does while drunk is another. With the notable exceptions of felony-murder and certain forms of accomplice liability, the common law does not allow the mens rea for one crime to substitute for the mens rea required for a second crime. Thus, in one famous case,³⁶ a thief broke a coin-operated gas meter to steal the coins and thereby caused a victim to be exposed to a frightful cloud of poison gas. Charged with both theft and the exposure of the victim to the gas, the court held that the intent to steal the

30. § 45-2-203.

31. *Egelhoff*, 518 U.S. at 41.

32. *Id.*

33. *Id.*

34. *See id.*

35. ARISTOTLE, NICOMACHEAN ETHICS bk. III, at 39–40 (David Ross ed., Oxford Univ. Press ed. 2009) (c. 384 B.C.E.).

36. *R v. Cunningham* [1957] 2 QB 396 (Eng.).

money could not substitute for the required mens rea concerning exposing the victim to the gas.³⁷ The mens rea for the two crimes had to be proven separately and independently.³⁸ The exceptions to this rule already noted are highly controversial precisely because they permit strict liability. For example, in felony-murder prosecutions, the mens rea for the underlying felony is sufficient to support a charge of murder if death results, even if the felonious defendant lacked the usual mens rea concerning death that is typically required to prove murder: purpose, knowledge, or recklessness under circumstances manifesting extreme indifference to the value of human life.³⁹

The influential Model Penal Code tries to have it both ways about intoxication.⁴⁰ While rejecting strict liability generally, the Code provides that a voluntarily intoxicated defendant may use evidence of such intoxication to negate purpose and knowledge but not to negate recklessness.⁴¹ The Code thus equates the culpability for becoming drunk with the conscious awareness of anything criminal that the agent might do while drunk. This “equation” permits the state to meet its burden of persuasion concerning recklessness without actually proving that the defendant was ever actually aware that getting drunk created a grave risk that the defendant would then commit the specific harm the statute prohibited. As an empirical matter, however, this equation is usually preposterous. An agent will not be consciously aware while becoming drunk that there is a substantial and unjustifiable risk that he or she will commit a particular crime when drunk, unless the person has a previous history of committing such harms. If prior history or other circumstances indicating previous conscious awareness exists, then the prosecution is capable of proving it and should be required to prove it. The prosecution should not be able to rely on what is, in effect, the conclusive presumption that becoming drunk demonstrates the same culpability as the actual conscious awareness of a substantial and unjustifiable risk that the defendant would commit the specific harm.

The Montana statute goes even further toward strict liability than the Model Penal Code, of course. One interpretation of the statute—rejected by Montana’s own Supreme Court but adopted by Justice Ginsburg—is that the intoxication provision simply works to redefine murder to include an

37. *Id.* at 397, 399–401.

38. *See id.* at 401.

39. See Stephen J. Morse, *Reason, Results, and Criminal Responsibility*, 2004 U. ILL. L. REV. 363, for a thorough critique of all forms of strict liability.

40. *See* MODEL PENAL CODE § 2.08(2) (AM. L. INST., Proposed Official Draft 1962).

41. *Id.*

objective *mens rea*: negligence.⁴² Ever since the Court's opinion in *Patterson v. New York*,⁴³ it has been clear that the states have the federal constitutional authority to effect such a redefinition, but this was not Montana's interpretation of its own law.⁴⁴ More important for my analysis, this redefinition undermines the standard view that culpability is hierarchically arrayed depending on the blameworthiness of the various mental states.⁴⁵ Our society's dominant morality simply does not accept, and with good reason, that negligent harmdoing is as blameworthy as committing the same harm purposely or with conscious awareness. The latter mental states indicate that the agent is consciously lacking in concern for the interests and well-being of an identifiable victim or class of victims, an attitude toward moral obligations that is more blameworthy than lack of awareness. Few except Oliver Wendell Holmes think that objective and subjective blameworthiness ought to be equated.⁴⁶ Characterizing a negligent killer as a murderer does violence to our ordinary notions of culpability and desert.

Consider Egelhoff's culpability again. First, assume that as the result of voluntary intoxication, he was in a mental state that qualifies as legal unconsciousness when he killed Pavola and Christenson. It is not morally unthinkable to condemn drinking oneself purposely or recklessly into a state of unconsciousness, but this behavior is not a crime *per se*. Criminal law theorists dispute the basis for the exculpatory effect of unconsciousness, but all agree that it does exculpate. Thus, if one believes Egelhoff's claim that he was legally unconscious, or to put it more accurately, if the prosecution were unable to prove beyond a reasonable doubt that he was legally conscious, then Egelhoff is not guilty of purposely or knowingly killing.⁴⁷ Moreover, there is no evidence that Egelhoff was consciously aware when he was drinking that he would become homicidal when drunk. Thus, he did not kill

42. *Montana v. Egelhoff*, 518 U.S. 37, 57–58 (1996) (Ginsburg, J., concurring).

43. 432 U.S. 197 (1977) (permitting New York to place on the defendant the burden of persuasion on the issue of "extreme emotional disturbance," New York's analogue to the provocation/passion doctrine, which traditionally reduces murder to manslaughter, and permitting the state largely to define, as it wishes, the elements of crime that the state must prove beyond a reasonable doubt).

44. *Egelhoff*, 518 U.S. at 41. Montana could, if it wished, and without constitutional hindrance, have defined the *mens rea* for murder as negligence. But such a definition would have been a similar abandonment of culpability and objectionable for precisely the same reasons I criticize the *Egelhoff* opinion. Once again, my goal is to consider society's response, not constitutional doctrine.

45. See Douglas Husak, *The Sequential Principle of Relative Culpability*, 1 LEGAL THEORY 493 (1995) (defending a qualified version of the claim that culpability is hierarchically arrayed depending on the relative blameworthiness of particular mental states).

46. See HOLMES, *supra* note 1, at 46–48.

47. The alternative theories that would support that result are unconsciousness negates the act requirement, negates *mens rea*, or provides an affirmative defense.

recklessly, even if one looks back to his earlier mental states to find culpability. Once again, Egelhoff might be fully responsible for becoming unconscious, but without proof of the mental states usually required, it is strict liability to hold him fully accountable for anything that he did while unconscious. He culpably caused the condition that would negate the prima facie case, but not with purpose, knowledge, or recklessness that he would be exonerated if charged with a homicide crime that requires one of these three subjective mental states.

Egelhoff is a dangerous agent, and it is undeniable that the State might have great difficulty proving beyond a reasonable doubt on these facts that he was legally conscious and thus guilty of purposely or knowingly killing. If he was legally conscious at the time of the killings, of course, the precision of the executions strongly suggests that he killed purposely or knowingly, even if one believes his claim that he did not remember the homicides. Without the crutch of strict liability, however, the State might be able to convict only for negligent homicide, typically graded as involuntary manslaughter, which carries a substantially shorter term of years than murder. But our fear of Egelhoff and revulsion at his deeds should not be allowed to promote injustice. The Constitution's requirement that in criminal cases the state must prove each element of the crime charged beyond a reasonable doubt almost always makes it more difficult for the prosecution to prove its most serious charge. Our society bears this risk because we believe that it is unacceptable to convict a legally innocent person. Concern with culpability thus almost always conflicts with concern for public safety. Weakening mens rea requirements and indirectly criminalizing noncriminal behavior or treating less serious criminal violations as major is an illiberal and surely ineffective remedy.

B. Mental Disorder, Intellectual Disability & Mens Rea

Doctrines that permit defendants to present mental disorder evidence to negate mens rea are often misleadingly termed "diminished capacity," mistakenly suggesting that these doctrines are kinds of mitigation or partial excuse.⁴⁸ They refer to a straightforward denial of a requisite element, akin to a denial that one acted or to a mistake of fact.⁴⁹ They are not a lesser form of the insanity defense. The failure to recognize this clear distinction often produces faulty reasoning about whether a defendant should be permitted to

48. Stephen J. Morse, *Undiminished Confusion in Diminished Capacity*, 75 J. CRIM. L. & CRIMINOLOGY 1, 6 (1984).

49. *Id.*

introduce mental disorder evidence to negate mens rea because courts wrongly believe the defendant is requesting the court to create a “mini” insanity defense.⁵⁰

Mental disorders relevant to mens rea most frequently produce disordered cognition, such as hallucinations or delusions, or untoward, sometimes strong desires, such as the desire to have sexual relations with minors or the desire to set fires or to take controlled substances. In other words, mental disorder produces crazy desires or crazy beliefs about reality, but it virtually never prevents a defendant from meeting the law’s criteria for intention, knowledge, conscious awareness, and other mens rea terms.⁵¹

Likewise, mental disorder seldom disables a defendant from having the *capacity* to form a mens rea.⁵² Modern inquiries into states of mind, both academic and judicial, seem obsessed with the vexed question of whether a defendant had the *capacity* to form mens rea, but either a defendant had the requisite mens rea or he did not. Mental disorder may in some cases demonstrate that the defendant did not form the mens rea at the time of the crime. Not having a mens rea or having a mental state inconsistent with the requisite mens rea does not mean, however, that someone was incapable of forming it. If an agent lacks the capacity to do something, it follows that the agent did not do it in fact, but if an agent had the capacity, it does not follow that he acted on it

Nevertheless, a defendant’s alleged capacity to form a mental state never provides better information than inquiring directly whether the mens rea was formed in fact, which is the ultimate legal question. Resolving questions about capacity requires a counterfactual inquiry that we lack the clinical and scientific resources to answer. When an expert testifies that a defendant lacked the capacity to form a mens rea, that opinion seldom has a clinical or scientific basis. It was precisely these types of difficulties that led California to bar testimony about the defendant’s capacity to form a mens rea, although it did permit testimony about whether the mental state was formed in fact.⁵³

50. See, e.g., *Bethea v. United States*, 365 A.2d 64, 84–85 (D.C. 1976). The Arizona case the Supreme Court considers in *Clark v. Arizona*, 548 U.S. 735 (2006), discussed below, makes precisely this error. *State v. Mott*, 931 P.2d 1046, 1050–51 (Ariz. 1997). Unfortunately, the Arizona Supreme Court recently perpetuated the confusion. See *State v. Malone*, 444 P.3d 733 (Ariz. 2019).

51. See Stephen J. Morse, *Crazy Reasons*, 10 J. CONTEMP. LEGAL ISSUES 189, 197–98 (1999).

52. Morse, *supra* note 48, at 42. Justice Breyer’s dissent in *Kahler v. Kansas* fully understood this point and its implications. 140 S. Ct. 1021, 1048–49 (2020) (Breyer, J., dissenting).

53. CAL. PENAL CODE § 28 (West 2021).

The examples of Daniel M’Naghten and Andrea Yates will help demonstrate that even the most delusional or hallucinating person can form the requisite mental state. M’Naghten delusionally believed that there was a conspiratorial Tory plot to kill him and formed a preemptive plan to kill the Tory Prime Minister, Robert Peel.⁵⁴ When he shot and killed Peel’s secretary, Drummond, believing the secretary was Peel, he surely intended to kill a person.⁵⁵ Likewise, Andrea Yates believed that unless she killed her children, they would become corrupt and would be tormented by Satan for eternity.⁵⁶ She therefore decided to kill her children.⁵⁷ She knew they were human beings and that human beings are killed by drowning. Ms. Yates surely intended to kill the five children when she drowned them in the bathtub.

Admittedly, on rare occasions, psychotic mentation is genuinely inconsistent with the formation of mens rea. In a well-known California case, the defendant, Wetmore, was caught in the victim’s apartment under conditions suggesting that he intended to steal the victim’s property.⁵⁸ Charged with burglary, the defendant claimed that he delusionally believed that the apartment and the property belonged to him.⁵⁹ If he told the truth, he did not intend to enter the apartment of *another* or to commit the felony of larceny, the elements of which include intentionally taking and carrying away the property of *another*.

Note that even if mental disorder does negate subjective mental states such as purpose, intention, knowledge, or conscious awareness of risk, it would never negate the objective negligence standard.⁶⁰ The person with mental disorder who is unaware of a risk that a reasonable person should be aware of is by definition unreasonable.⁶¹ Even the Model Penal Code, which individualizes the negligence inquiry somewhat by requiring the decision maker to consider the behavior of a “reasonable person . . . *in the actor’s*

54. RICHARD MORAN, KNOWING RIGHT FROM WRONG: THE INSANITY DEFENSE OF DANIEL MCNAUGHTAN 10 (1981).

55. M’Naghten’s Case (1843) 8 Eng. Rep. 718, 719.

56. Phillip J. Resnick, *The Andrea Yates Case: Insanity on Trial*, 55 CLEV. ST. L. REV. 147, 149 (2007); see also Deborah W. Denno, *Who Is Andrea Yates? A Short Story About Insanity*, 10 DUKE J. GENDER L. & POL’Y 1 (2003) (providing a thorough account of the entire background and an analysis of the expert testimony).

57. Resnick, *supra* note 56, at 149; see also Denno, *supra* note 56, at 2.

58. *People v. Wetmore*, 583 P.2d 1308, 1310–11 (Cal. 1978).

59. *Id.* at 1310.

60. *E.g.*, *Bashi v. Wodarz*, 53 Cal. Rptr. 2d 635, 641 (Ct. App. 1996) (holding that sudden mental illness did not, as a matter of law, preclude liability for negligence).

61. See Grant H. Morris, *Requiring Sound Judgments of Unsound Minds: Tort Liability and the Limits of Therapeutic Jurisprudence*, 47 SMUL. REV. 1837, 1838–39 (1994).

situation” before making a finding of negligence,⁶² would not go so far as to consider mental disorder—irrational behavior—as part of the “situation.” The Model Penal Code never gives a clear definition of the “situation,” but it does make clear that it wishes to avoid complete subjectification of the reasonable person standard.⁶³ To assess reasonableness from the standpoint of the “reasonable irrational” person would deprive the negligence standard of all objectivity.⁶⁴

In other cases, mental disorder may not necessarily be inconsistent with formation of mens rea, but evidence of disorder may help bolster the defendant’s claim that he did not form it. For example, suppose a psychotically disorganized person gets lost in an empty part of town on a cold winter’s night and cannot find his way home. To escape the cold, he breaks into a building, is caught, and is charged with burglary on the theory that he intended to steal. In this case, he is fully capable of forming the intent to steal, but his mental disorder helps explain why he broke in simply to keep warm.

Again, the crucial issue is to determine the defendant’s actual mental state and to compare that mental state to the mental state required by the crime charged. Of course, the lurking problem is that it is sometimes very difficult to determine a defendant’s mental state at the time of the crime, but these difficulties arise in all retrospective mental state evaluations, and not just in cases involving mental disorder. Nevertheless, one of the primary skills we have evolved over the last 100,000 years is the ability to discern the intentions of other humans. Like Holmes’s dog, jurors are quite good at recognizing the difference between a kick and a stumble.

In sum, the mens rea issue is entirely distinct from the legal insanity issue, even if precisely the same evidence would be relevant to adjudicating both claims. People with mental disorders are not automatons; rather, they are agents who act for reasons. Their reasons may be motivated by distorted perceptions and beliefs, but they do form intentions and have knowledge of what they are doing in the narrow, most literal sense. Thus, it is very uncommon for mental disorder to negate all mens rea, even if the defendant is profoundly delusional. In some cases, the same evidence that a defendant

62. MODEL PENAL CODE § 2.02(2)(d) (AM. L. INST., Proposed Official Draft 1962) (emphasis added).

63. MODEL PENAL CODE AND COMMENTARIES § 2.02, at 242 (AM. L. INST. 1985).

64. Failure to recognize this point and the incorrect belief that the Model Penal Code adopts nearly complete subjectification for negligence are major analytic reasons that the primary contemporary proposal to abolish the insanity defense fails. *See generally* Christopher Slobogin, *An End to Insanity: Recasting the Role of Mental Disability in Criminal Cases*, 86 VA. L. REV. 1199, 1239 (2000) (arguing that the Model Penal Code’s definition of negligence is subjective).

was delusional may both negate mens rea and support a finding of legal insanity, but the questions being answered by the evidence are different.⁶⁵

With this background in mind, let us turn to *Clark*. In the early morning of June 21, 2000, Eric Clark, a seventeen-year-old resident of Flagstaff, Arizona, was riding around in his pickup truck blaring loud music.⁶⁶ Responding to complaints about the noise, Officer Jeffrey Moritz, who was in uniform, turned on the emergency lights and siren of his marked patrol car and pulled Clark over.⁶⁷ Moritz left the patrol car and told Clark to remain where he was.⁶⁸ Less than a minute later, Clark shot and killed Moritz.⁶⁹

Clark was charged with intentionally killing a police officer, knowing that the officer was acting in the line of duty.⁷⁰ He did not contest the shooting and death, but he claimed that as a result of paranoid schizophrenic delusions, he lacked the required mens rea for the crime charged (the intent to kill a person and the knowledge that the victim was a police officer).⁷¹

Substantial evidence, including Clark's statements to classmates a few weeks earlier that he wanted to shoot police officers, suggested that Clark knew Moritz was a police officer and that he had planned just such a shooting.⁷² He had even arguably lured Officer Moritz by driving his truck with its radio blaring in a residential area.⁷³ On the other hand, Clark presented testimony from family, classmates, and school officials about his bizarre behavior during the preceding year, including rigging his bedroom with fishing line, beads, and chimes to warn him of intruders, and keeping a bird in his car to warn him of airborne poison.⁷⁴ These actions were plausibly a result of his paranoid delusions. Indeed, "[t]here was lay and expert testimony that Clark thought Flagstaff was populated with 'aliens,'" including some that were impersonating police officers, that the aliens were trying to kill him, and that only bullets could stop the aliens.⁷⁵ The defense

65. See *Clark v. Arizona*, 548 U.S. 735 (2006). For example, if Clark really believed he was killing an alien impersonating a police officer, he lacked knowledge that he was killing a police officer, thus negating the mens rea of knowledge. For the same reason, he did not know what he was doing or that it was wrong.

66. *Id.* at 743.

67. *Id.*

68. *Id.*

69. *Id.*

70. *Id.*

71. *Id.* at 743–44.

72. *Id.*

73. *Id.*

74. *Id.* at 745.

75. *Id.*

expert also testified that Clark may have turned the radio up to drown out auditory hallucinations.⁷⁶

The operative Arizona rule concerning the admission of evidence of mental disorder to negate mens rea was based on an Arizona Supreme Court decision, *State v. Mott*, which held that psychiatric testimony was inadmissible to negate specific intent and that evidence of mental disorder was not admissible to negate any mens rea element.⁷⁷

At a bench trial, the judge permitted introduction of all the lay and expert testimony about Clark's mental disorder at the time of the crime.⁷⁸ The judge specifically found that Clark was suffering from paranoid schizophrenia and had distorted perceptions at the time of the crime.⁷⁹ He ultimately ruled, however, that *Mott* barred him from using mental disorder evidence to consider mens rea and that he could consider this evidence only to decide the issue of legal insanity.⁸⁰ He found Clark guilty of first-degree murder.⁸¹

Clark appealed on the grounds that the *Mott* rule violated procedural due process.⁸² The Arizona Court of Appeals affirmed the conviction, the Arizona Supreme Court denied review, and the United States Supreme Court granted certiorari to decide "whether due process prohibits Arizona . . . from excluding evidence of mental illness and incapacity due to mental illness to rebut evidence of the requisite criminal intent."⁸³ In addressing this issue, the Court's majority got it quite wrong, not only confounding mens rea and legal insanity, but forcing psychiatric and psychological evidence into an arbitrary system of classification that is unworkable and that had never been proposed or argued by Arizona.

Although the extent of the *Mott* rule's evidentiary exclusion is not entirely clear, it prevents defendants from introducing substantial, relevant, and reliable mental health expert testimony concerning whether or not the requisite mens rea was formed in fact. Recall Clark's factual claim that, as a result of delusions, he actually believed Officer Moritz was a dangerous "alien" impersonating a police officer.⁸⁴ If this were true, then Clark would

76. *Id.* at 782.

77. 931 P.2d 1046, 1051 (Ariz. 1997). The Arizona Supreme Court mischaracterized the use of battered woman syndrome evidence offered to negate mens rea as a "diminished capacity" defense. *Id.* at 1050. The court then held, consistent with its mistaken premise, that the Arizona legislature had implicitly rejected that defense when it refused to recognize generalized diminished capacity and ruled that the evidence could be excluded. *Id.* at 1050, 1055.

78. *Clark*, 548 U.S. at 745.

79. *Id.* at 746.

80. *Id.* at 745-46.

81. *Id.*

82. *Id.* at 746.

83. *Id.* at 746-47.

84. *Id.* at 745.

not be guilty of intentionally killing a person, let alone guilty of knowingly killing a police officer acting in the line of duty. Clark was functionally prevented by the trial judge's reliance on *Mott* from using relevant and reliable testimony to cast a reasonable doubt about the mens rea for the crime charged.

1. The Tripartite Evidence Construction

In order to insulate the *Mott* rule from Clark's due process claim, the Supreme Court majority mischaracterized Arizona's rule. Instead of holding that *Mott* excluded expert mental health testimony on the issue of mens rea, the Court ruled that *Mott* actually applied only to two of the three types of distinguishable mens rea evidence: observation evidence, mental-disease evidence, and capacity evidence.⁸⁵ The Court then upheld the rule against a due process attack simply by announcing that the Arizona courts could not possibly have intended it to apply to the first category.⁸⁶

The majority defined the first category, "observation evidence," "in the everyday sense, [as] testimony from those who observed what [a defendant] did and heard what he said; this category would also include testimony that an expert witness might give about [the specific defendant's] tendency to think in a certain way and his behavioral characteristics."⁸⁷ The Court pronounced that this first category of mens rea evidence was admissible despite *Mott* and could be presented by either lay or expert witnesses.⁸⁸

The second category, "mental-disease evidence," was defined as "opinion testimony that [a defendant] suffered from a mental disease with features described by the witness."⁸⁹ That is, such testimony provides general information about the mental disorder from which the defendant allegedly suffers. The Court pronounced this kind of expert evidence inadmissible under *Mott*.⁹⁰ The majority defined the third category, "capacity evidence," as evidence "about a defendant's capacity for cognition and moral judgment

85. *Id.* at 757–59.

86. *Id.* at 760.

87. *Id.* at 757.

88. *Id.* at 760–61. It is interesting that the majority constructs these categories of mens rea evidence without being explicit about whether they are meant to describe the actual limits of *Mott* (an odd thing for a federal court to be doing when the rule at issue was invented by a state court) or the constitutional limit of *Mott*'s reach. It must be the latter, but the majority's reluctance to admit it is puzzling.

89. *Id.* at 758.

90. *Id.* at 760.

(and ultimately also his capacity to form *mens rea*).⁹¹ This, too, is expert opinion evidence, and again the Court upheld *Mott*'s exclusion of it.⁹²

In upholding the constitutionality of *Mott*'s bar to mens rea evidence in categories two and three, the *Clark* majority mistakenly refers to barring these types of evidence on the legal insanity issue, not mens rea: "Thus, only opinion testimony going to mental defect or disease, and its effect on the cognitive or moral capacities on which *sanity* depends under the Arizona rule, is restricted."⁹³ Is this astonishing error typographical, or a deeper reflection of the majority's own confusion between these two very different principles?⁹⁴

The majority's unprecedented tripartite construction fails to do the theoretical work necessary to draw a sensible line between which types of expert mental disorder evidence states may and may not exclude. This classification was not part of Arizona law (nor any state law I know about) and cannot be found in any Supreme Court precedent. If that were not enough, neither of the parties or amici ever suggested such a construction in their briefs or at oral argument.

Justice Kennedy's dissent, and to some extent Justice Breyer's concurrence, also recognized that the majority's tripartite classification quickly breaks down in practice and therefore ends up being an unworkable solution to this constitutional problem.⁹⁵ There are clear, core cases of each type, but testimony rarely comes so neatly packaged. For example, sound empirical evidence about the characteristics of people suffering from a particular mental disorder is based on observation and is factual. In many cases, such evidence would help a finder of fact understand the behavior of a defendant who suffers from that disorder, even though the evidence comes from the observations of others. Much of the inferential expert evidence Justice Souter believes *Mott* may constitutionally exclude is actually observational evidence that could not have been excluded had the observations been about the defendant himself. Would experts be allowed to testify about their own observations of the defendant, but then not be permitted to testify about the features of a recognized diagnostic category that help explain those observations?

91. *Id.* at 758.

92. *Id.* at 760.

93. *Id.* (emphasis added).

94. This is not an isolated example of the majority conflating insanity and mens rea. See *infra* text accompanying notes 96–98.

95. *Clark*, 548 U.S. at 781–84 (Kennedy, J., dissenting); *id.* at 779–80 (Breyer, J., concurring).

All clinical judgments in medicine, psychiatry, and clinical psychology are, by their very nature, informed by the clinician's observation of the particular patient being seen, by the accumulated wisdom of observations of other patients by that clinician and others, and by findings from empirical studies. The Court's fictitious categories of evidence not only bleed into one another, but they also seem peculiarly unable to do the constitutional work the Court asks of them.

The Court observed that the presumption of sanity is universal in some form but confusingly mischaracterized it as "a presumption that a defendant has the capacity to form the *mens rea* necessary for a verdict of guilt and the consequent criminal responsibility."⁹⁶ Once again, the Court confused *mens rea* with insanity, this time in the heart of the opinion explaining why Arizona may constitutionally prevent a defendant from introducing category two and three evidence to negate *mens rea*.⁹⁷ Indeed, the whole presumption-of-sanity discussion in this section of the opinion is irrelevant to Clark's *mens rea* claim.

The Court not only blurred the distinction between *mens rea* and insanity but also made wholly inconsistent observations about the relationship between the two. It rejected the argument that *mens rea* and insanity are "entirely distinguishable," yet obscurely noted that insanity "trumps" *mens rea*, suggesting that they are not functionally distinguishable because the former subsumes the latter when both are claimed.⁹⁸

The majority correctly acknowledged that evidence of the defendant's state of mind at the time of the crime might indicate the defendant's actual mental state and the presence of an enduring incapacity to form the requisite *mens rea*.⁹⁹ Despite that acknowledgment and the majority's recognition of a defendant's constitutional right to present evidence that negates an element, the Court nonetheless concluded that Arizona could constitutionally deprive Clark of the right to present relevant *mens rea* evidence. The Court claimed that Arizona could permissibly "channel" mental disorder evidence solely into the insanity issue because Arizona had legitimate state interests in such channeling.¹⁰⁰

96. *Id.* at 766 (majority opinion).

97. *Id.*

98. *Id.* at 768 n.38.

99. *Id.* at 767.

100. *Id.* at 779.

2. The Channeling Argument

As *Montana v. Egelhoff* makes clear, states may of course preclude relevant defense evidence—even evidence that rebuts an element of the offense—if the state has a legitimate purpose in excluding the evidence.¹⁰¹ But the Court’s analysis of Arizona’s “legitimate purposes” in channeling some forms of expert mental state evidence into insanity goes radically wrong.

The Court’s first identified “legitimate reason” is yet another example of its conflation of mens rea and legal insanity. The Court reaffirmed Arizona’s authority to define legal insanity as it wishes and to place the burden of persuasion for this defense on the defendant.¹⁰² Consequently, the Court reasoned that if Arizona “is to have this authority in practice as well as in theory, it must be able to deny a defendant the opportunity to displace the presumption of sanity more easily when addressing a different issue.”¹⁰³ This is a non sequitur that is based on confusion about the presumption of sanity and “capacity” that we addressed above.¹⁰⁴ There is no presumption of sanity applicable to the mens rea elements. Even if there were, the presumption must be rebuttable, and in a case involving severe mental disorder. Introducing expert testimony is a crucial method of rebuttal.

Mental disorder and capacity evidence used to negate mens rea could result in acquittal simply by presenting a reasonable doubt about mens rea, whereas the same type of evidence used to prove legal insanity would succeed in Arizona only if the defendant convinced the finder of fact of his insanity by a preponderance of the evidence. Such a difference is not a “displacement” of the presumption of sanity, however. It is simply a logical consequence of the interaction between the structure of criminal culpability and *Winship*’s constitutional requirements.¹⁰⁵ Criminal responsibility can be avoided either by negating an element of the crime charged or by establishing an affirmative defense. Permitting an affirmative defense does not remove the State’s obligation to prove the elements beyond a reasonable doubt. Clark’s claim does not undermine Arizona’s right to define legal insanity as it wishes.

The Court’s second “legitimate purpose” is Arizona’s desire “to avoid a second avenue for exploring capacity, less stringent for a defendant.”¹⁰⁶

101. 518 U.S. 37 (1996); *see supra* text accompanying note 44.

102. *Clark*, 548 U.S. at 771–72.

103. *Id.* at 771.

104. *See supra* text accompanying notes 96–98.

105. *In re Winship*, 397 U.S. 358, 368 (1970) (holding that the Constitution only allows a criminal defendant to be convicted when all facts necessary to constitute a crime are proven beyond a reasonable doubt).

106. *Clark*, 548 U.S. at 770, 772.

Contrary to the Court's assertion, permitting a jury to use mental-disease or capacity evidence to decide if there is a reasonable doubt about mens rea is *not* "in functional terms . . . analogous to allowing jurors to decide upon some degree of diminished capacity to obey the law . . . that would prevail as a standalone defense."¹⁰⁷ Negation of mens rea is not an independent "defense" of "diminished capacity." Mens rea negation is simply a straightforward denial of the prima facie case. A state undeniably has the authority to reject a partial responsibility mitigating doctrine, but this was not what Clark was requesting.

The Court then articulated a third legitimate purpose: Arizona has made a determination that mental state evidence is too unreliable to be the basis of a complete defense under which a defendant is entirely acquitted and unconditionally freed.¹⁰⁸ Now, at last, we are at the heart of Arizona's real concerns. In fact, the Court identified three separate concerns: (1) "the controversial character of some categories of mental disease," (2) "the potential of mental-disease evidence to mislead," and (3) "the danger of according greater certainty to capacity evidence than experts claim for it."¹⁰⁹ The Court's general conclusion about the second and third risks was that shifting the burden of persuasion to the defendant by channeling the evidence into the insanity issue would reduce the risk that misleading evidence would lead to incorrect verdicts.¹¹⁰

The Court is certainly correct and has noted on many occasions that there is great debate about the concept of and criteria for mental disorders. As a result, caution is warranted "in treating psychological classifications as predicates for *excusing* otherwise criminal conduct."¹¹¹ This is true enough, but the same argument applies to any use of diagnostic information, which is routinely admitted in a wide array of civil and criminal law contexts, including the insanity defense. The Court's argument proves too much.

Moreover, the defendant is not seeking to excuse his conduct. This might be a valid reason for channeling in the other direction: allowing mental disorder evidence for mens rea and disallowing it for insanity. It is not a justification for channeling in the *Mott* direction, which *allows* excusing

107. *Id.* at 773.

108. *Id.* at 776–78.

109. *Id.* at 774.

110. *Id.* at 778.

111. *Id.* at 774 (emphasis added). I am a long-time critic of the use of diagnostic information in civil or criminal cases. Stephen J. Morse, *Crazy Behavior, Morals, and Science: An Analysis of Mental Health Law*, 51 S. CAL. L. REV. 527, 604–15 (1978); Stephen J. Morse, *Failed Explanations and Criminal Responsibility: Experts and the Unconscious*, 68 VA. L. REV. 971, 1055–70 (1982).

evidence but *disallows* evidence that would exonerate because it defeats the prima facie case.¹¹²

The second risk—that mental-disease, i.e., general diagnostic, evidence may lead to the incorrect conclusion that the defendant lacks capacity to form mens rea when in fact he possessed mens rea—is true to a degree. Avoiding such a risk is no doubt a legitimate state interest. There is great behavioral heterogeneity within psychiatric diagnostic categories and imperfect fit, as the American Psychiatric Association recognizes, between those categories and legal questions.¹¹³ Testifying experts, alas, do not always confine themselves to providing rigorously confirmed evidence about the characteristics of people like the defendant, and instead fall into the trap of using diagnostic terms that do not inform the legal issues. It is all too easy for professionals and lay people alike to make the mistake of begging legal questions based on a psychiatric or psychological diagnosis.

Indeed, the Court pointed out that the testifying experts in *Clark* made this error themselves while testifying about legal insanity: they agreed on the diagnosis of schizophrenia but disagreed about Clark's cognitive and moral capacity.¹¹⁴ Given the dangers of mental-disease evidence to mislead, the Court concluded (and I partially agree) that it is reasonable for a state to decide to channel the evidence to the insanity defense on which the defendant can be assigned the burden of persuasion.¹¹⁵

This justification for complete channeling of expert evidence to the insanity issue nevertheless again promotes injustice. As with some of the Court's other justifications, there is no reason a state's skepticism about psychiatric evidence should begin and end with the criminal law. States use diagnostic information in a host of other legal contexts. If anything, the law should be more forgiving when criminal blame and punishment are at stake. Moreover, the problem arises less from the inherent tendency of diagnostic information to mislead than from confusion about the nature of the relation of such categories to a legal conclusion. Too often people wrongly believe that if a mental disorder played a causal role, the behavior is akin to a mechanism, and the defendant is therefore not responsible. Further problems arise from the failures of the trial process adequately to cabin the experts.

Note that the experts in *Clark* disagreed about the defendant's cognitive and moral capacities—that is, about whether he was legally insane. But insanity is a legal question to be resolved by a lay jury or judge. Why should

112. *State v. Mott*, 931 P.2d 1046, 539–46 (Ariz. 1997).

113. AM. PSYCHIATRIC ASS'N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 25 (5th ed. 2013) [hereinafter DSM-5].

114. *Clark*, 548 U.S. at 758–59.

115. *Id.* at 774.

we expect mental health experts to agree about this ultimate legal question about which they have no more expertise than lay jurors or judges? Indeed, it is for precisely this reason that expert witnesses in federal criminal trials are not permitted to offer ultimate legal conclusions about whether a defendant was legally insane.¹¹⁶ Simple and sensible evidentiary rules like this would diminish the misleading tendencies of diagnostic information without the need for draconian rules that prevent defendants from defending themselves with relevant evidence.

The Court's third and last argument justifying Arizona's channeling rule considered the dangers that capacity evidence allegedly presents. The Court pointed out that opinions about the capacity for moral cognition or to form mens rea are inferential judgments "fraught with multiple perils," including accurately determining the defendant's mental state at the time of the crime and properly understanding the differences between psychological and legal judgments about capacity.¹¹⁷ Moreover, testimony about the defendant's capacity to form mens rea is essentially "ultimate legal issue" testimony about which mental health experts have no special expertise.¹¹⁸ According to the Court, there is a real risk that the expert's judgment about capacity will have an apparent authority that honest mental health professionals do not claim to have.¹¹⁹ States may reasonably address these dangers by channeling capacity evidence to the insanity issue and placing the burden of persuasion on the defendant.

The Court's basic critique of capacity evidence is sound. Evidence about a defendant's capacity to form mens rea is problematic and often lacks a solid clinical or scientific foundation. The criminal law would be better off if capacity evidence were strictly limited or even prohibited. The Court's argument nevertheless proves too much because capacity evidence is admissible almost everywhere to address every other question in criminal and civil law to which it may be relevant.¹²⁰ Why permit its limitation here, when so much is at stake for the criminal defendant? Furthermore, permitting an expert to give an opinion about legal insanity—which experts are permitted to do almost everywhere—is *a fortiori* ultimate issue testimony, which the Court rejects concerning the mens rea issue.

116. FED. R. EVID. 704(b).

117. *Clark*, 548 U.S. at 776–77.

118. *Id.* at 777–78 (quoting Insanity Work Def. Grp., *American Psychiatric Association Statement on the Insanity Defense*, 140 AM. J. PSYCHIATRY 681, 686 (1983)).

119. *Id.* at 778.

120. California prohibits experts in criminal trials from offering an opinion about whether the defendant had the capacity to form mens rea. See CAL. PENAL CODE §§ 29.2, 29.4(b) (West 2021).

Moreover, the Court's capacity argument was marred by providing all its examples from the context of legal insanity, yet again blurring the two doctrines. The ultimate issue of legal insanity is considerably less factual than the ultimate issue of mens rea. The former does depend on a factual understanding of the defendant's mental state at the time of the crime, but the finder of fact ultimately must make a normative moral judgment that is not straightforwardly factual because the borders of all insanity tests are fuzzy and open to interpretation. In contrast, whether a defendant formed a requisite mens rea is, with few exceptions,¹²¹ a purely factual question. As Justice Kennedy said about Clark in his dissent, "Either Clark knew he was killing a police officer or he did not."¹²² The question is not about moral responsibility but about empirical fact; therefore, the expert is less likely to mislead about mens rea than about legal insanity. Again, this would argue for channeling the evidence in the opposite direction.

The central problem with the majority's channeling argument is one of basic fairness. When a citizen is being threatened with the most awesome exercise of state power—criminal blame and punishment—it seems that we should be most permissive in allowing that citizen to defend himself with the same relevant and reliable evidence allowed without limitation in all other legal contexts. What is it about the criminal trial that drives the Court to tolerate defense handicaps it would not tolerate in any other arena?

It is instructive to compare the Court's approach to psychiatric evidence in another case, *Barefoot v. Estelle*, which involved a due process challenge to the admission of a prosecution expert's psychiatric opinion about a defendant's future dangerousness.¹²³ The opinion was elicited in the death penalty phase of a capital trial and was based entirely on the psychiatrist's responses to the prosecutor's hypothetical questions.¹²⁴ *Barefoot* claimed, with the support of all the relevant mental health organizations as amici, that clinical predictions by mental health professionals of a defendant's future dangerous conduct were so inaccurate that they would inevitably lead to erroneous sentences.¹²⁵ Although capital punishment was at stake, the Court upheld the admission of such predictions and ruled that the deficiencies of the testimony went simply to its weight and could be addressed by cross-examination.¹²⁶

121. For example, the "premeditation" standard that in many jurisdictions makes an intentional killing a first-degree murder often involves some degree of normative evaluation. Hate motivation for hate crimes is another example.

122. *Clark*, 548 U.S. at 790 (Kennedy, J., dissenting).

123. 463 U.S. 880 (1983).

124. *Id.* at 884.

125. *Id.* at 898–901.

126. *Id.* at 898, 903.

All the same arguments the Court mounts in *Clark* against mental-disease and capacity evidence apply *a fortiori* to predictions of dangerousness, and there is no reason the same remedy that saved the potentially misleading evidence in *Barefoot*—vigorous cross-examination to expose its defects—could not apply equally in *Clark*. How can it be fair to let the state present problematic mental health evidence to support imposition of capital punishment but deny the defendant the right to use similar evidence to defend himself against a charge that he even committed a crime?

The Court in *Clark* never satisfactorily addresses this basic issue of fairness or the extraordinary degree to which these kinds of channeling rules compromise the right of citizens to demonstrate their innocence.¹²⁷ *Clark* did not have a fair chance to cast reasonable doubt on the mens rea necessary to convict him of first-degree murder.

3. Exclusion of Mental Abnormality Evidence To Negate Mens Rea Is Wrong

Whether states may constitutionally abolish mens rea has been an unresolved question of constitutional law ever since the emergence of the regulatory state and the concomitant growth and acceptability of strict liability crimes.¹²⁸ It has been clear at least since *Morissette v. United States*, however, that mens rea is essential to the criminal law, and that states and Congress bear a heavy burden of justifying departures from it.¹²⁹ Nevertheless, the Supreme Court has never crossed the line from “essential” to “constitutionally required,” even for common law crimes at the core of the criminal law whose mens rea elements predate the Constitution. No legislature has ever been so bold as to purport to convert a serious common law crime into a strict liability crime. Consequently, this difficult constitutional question has lingered at the edges of justiciability, drawing the considerable attention of academics but not courts.¹³⁰ Nevertheless, in an era

127. See *Clark*, 548 U.S. at 770–71.

128. See, e.g., *United States v. Dotterweich*, 320 U.S. 277 (1943) (upholding strict liability and punishment for up to a year for shipping misbranded pharmaceuticals); *United States v. Balint*, 258 U.S. 250 (1922) (upholding strict liability and punishment for up to five years for selling controlled substances without the order form required by the Narcotic Act of 1914).

129. *Morissette v. United States*, 342 U.S. 246, 252, 262–63 (1952).

130. Compare Herbert L. Packer, *Mens Rea and the Supreme Court*, 1962 SUP. CT. REV. 107, 126–27 (suggesting constitutional limits to the legislative abolition of mens rea), with Louis D. Bilionis, *Process, the Constitution, and Substantive Criminal Law*, 96 MICH. L. REV. 1269, 1278–79 (1998) (arguing that individual blameworthiness has been “mythicized” as traditional criminal law). Complicating the debate is the impact of punishment theory. Some commentators have

in which criminal justice reform aims at making mens rea requirements more robust, the opposite trend in the mental health context is anomalous and wrong.

That mens rea must be part of any constitutionally legitimate criminal system is not so much an historical conclusion as a moral one. It is immoral for the state to punish people whose mental disorders prevented them from forming mens rea. Mens rea is a fundamental feature of our moral practices. For those with and without mental disorder, mens rea expresses the defendant's attitudes towards the rights and interests of fellow citizens. Different culpable mental states justify different levels of punishment consistent with different degrees of desert. Without mens rea, no blame and punishment would be justified at all, at least for core crimes.

Mens rea is so crucial to fair ascriptions of blame and imposition of punishment that we should be wary of attempts to dilute it by redefinitions of elements and affirmative offenses or by evidentiary rules. Although the criminal law can never guarantee perfect justice and must often balance competing moral and practical considerations, the risk of error should seldom be shifted to the defendant unless there are supremely good reasons for doing so. The presumption should be against such redefinitions and evidentiary rules that undermine the values *Winship* protects.¹³¹

C. *The Affirmative Defense of Legal Insanity*

In the course of upholding the constitutionality of Arizona's extremely narrow legal insanity test in *Clark*, the Supreme Court noted that it had never ruled whether an affirmative defense of legal insanity is required by due process.¹³² A few older state court opinions held abolition of the defense unconstitutional as a violation of fundamental fairness and the right to a jury trial.¹³³ In the later part of the twentieth century, five states abolished the insanity defense, and in four of the states the state supreme court upheld the

argued that a retributionist might, in some circumstances, care less about mens rea than a rehabilitationist. See, e.g., Kenneth W. Simons, *When Is Strict Criminal Liability Just?*, 87 J. CRIM. L. & CRIMINOLOGY 1075, 1130–31 (1997). Though this detour is beyond the scope of this article, I must mention here that though I count myself as essentially in the retributionist camp, I do not share the idea that mens rea is separable from blameworthiness. On the contrary, the very idea of blameworthiness likely has its roots in our evolutionary ability to recognize intentionality.

131. *In re Winship*, 397 U.S. 358, 363–64 (1970).

132. *Clark*, 548 U.S. at 752 n.20

133. *Sinclair v. State*, 132 So. 581, 582 (Miss. 1931); *State v. Strasburg*, 110 P. 1020, 1024 (Wash. 1910).

abolition.¹³⁴ In the fifth state, the Nevada Supreme Court held that abolition violated due process under both the state and federal constitutions.¹³⁵

Finally, in *Kahler v. Kansas*, the Supreme Court held that the Constitution does not require a state to provide an independent, affirmative defense of insanity.¹³⁶ It was sufficient, the majority held, if the jurisdiction permitted the defendant to introduce evidence of mental abnormality to negate the mens rea for the crime charged, a practice known as the mens rea alternative to the insanity defense.¹³⁷ The Court opined that this rule in effect introduced a cognitive test for legal insanity to defeat the prosecution's prima facie case.¹³⁸ In so doing, the Court and states employing the mens rea alternative implicitly conceded that mental disorder is sometimes relevant to fair blame and punishment. I have long thought that the Constitution does not require an independent affirmative defense of legal insanity if an alternative that would do equal justice could be found. The Court thought it had succeeded in finding a just alternative to an affirmative defense of legal insanity, but this alternative is a practical failure that denies justice.

The moral and legal necessity of some form of an insanity defense or a reasonable alternative is easily explained. The capacity for rationality and, more controversially, for self-control are the touchstones of responsibility, and their lack explains almost all the partial and complete affirmative defenses.¹³⁹ In some severe cases, mental abnormality markedly diminishes those capacities, and thus the defendant is not a morally responsible agent who deserves blame and punishment. It is for similar reasons based on developmental immaturity and intellectual disability (itself a form of mental abnormality) that some juveniles and some defendants with intellectual disability do not deserve the most severe punishments even if they retain some degree of criminal responsibility. There was an unifying and

134. *State v. Bethel*, 66 P.3d 840, 851–52 (Kan. 2003); *State v. Herrera*, 895 P.2d 359, 367–68 (Utah 1995); *State v. Searcy*, 798 P.2d 914, 919 (Idaho 1990); *State v. Korell*, 690 P.2d 992, 999–1000 (Mont. 1984). The abolition by four states was the first time since the institution of legal insanity in the 1300s that an Anglo-American jurisdiction did not have some form of the defense.

135. *Finger v. State*, 27 P.3d 66, 68 (Nev. 2001).

136. 140 S. Ct. 1021 (2020).

137. *Id.* at 1037. The Court uses the term “approach” rather than “alternative,” but it is an alternative and the latter locution is more faithful to the structure of the moral argument.

138. *Id.* at 1025–26.

139. Duress is the obvious counterexample. There may be nothing wrong with the defendant's capacities for rationality and self-control, but we simply cannot expect people to do the right thing in these cases. It is too much to ask of reasonable, law-abiding people when they are sufficiently threatened. It is such reasoning that leads some to think that duress is best classified as a justification. See R. JAY WALLACE, *RESPONSIBILITY AND THE MORAL SENTIMENTS* (1994); see also Peter Westen & James Mangiafico, *The Criminal Defense of Duress: A Justification, Not an Excuse—and Why It Matters*, 6 *BUFF. CRIM. L. REV.* 833 (2003).

inconclusive dispute between the *Kahler* majority and dissent about whether the history supported the centuries-old existence of some form of an independent insanity defense. It was unedifying and inconclusive because until well into the twentieth century, courts, legal commentators, and treatise writers often confused mental disorder as negating the elements with mental disorder as the basis of an affirmative defense. That confusion was apparent in both the *Clark* and *Kahler* majorities. Thus, no clear answer could be expected from the historical sources. What is clear, however, is that in one form or another, mental disorder has had exculpatory force since the Late Middle Ages in English law and since the founding in U.S. law.¹⁴⁰

A similar baseline principle explains the many competence doctrines employed in the criminal justice process. The Court has long recognized that at every stage justice demands that some people with severe mental abnormalities must be treated differently from those without substantial mental impairment because some impaired defendants are incapable of reason and understanding in a specific context. Competence to stand trial,¹⁴¹ competence to plead guilty and to waive counsel,¹⁴² competence to represent oneself,¹⁴³ and competence to be executed¹⁴⁴ are all examples in which the Constitution requires such special treatment. It is unfair and offensive to the dignity of criminal justice to treat people without understanding as if their understanding was unimpaired. Evidence of mental disorder is routinely introduced in all these contexts to determine if the defendant must be accorded special treatment.

The negation of mens rea and the affirmative defense of legal insanity are different claims that preclude criminal liability by different means. The post-verdict consequences are also different. The former leads to outright acquittal; the latter results in some form of involuntary civil commitment. Although in some cases the same mental disorder evidence may be used to prove the two different claims, they are not equivalent.

The primary reason that permitting a defendant to introduce evidence of mental disorder to negate mens rea cannot justly replace the affirmative defense of legal insanity is that the mens rea alternative is based on a mistaken view of how severe mental disorder affects human behavior that was

140. Brief of 290 Criminal Law and Mental Health Law Professors as Amici Curiae Supporting Petitioner's Request for Reversal and Remand at 4–5, *Kahler v. Kansas*, 140 S. Ct. 1021 (2020) (No. 18-6135).

141. *Drope v. Missouri*, 420 U.S. 162 (1975).

142. *Godinez v. Moran*, 509 U.S. 389 (1993).

143. *Indiana v. Edwards*, 554 U.S. 164 (2008).

144. *Ford v. Wainwright*, 477 U.S. 399 (1986); *Panetti v. Quarterman*, 551 U.S. 930 (2007); *Madison v. Alabama*, 139 S. Ct. 718 (2019).

previously discussed.¹⁴⁵ In virtually all cases, mental disorders, even severe disorders marked by psychotic symptoms such as delusions and hallucinations, do *not* negate mens rea. It is difficult to prove a negative, but cases in which most or all mens rea is negated are rare to the vanishing point, especially cases of serious crime. Rather, mental disorder affects a person's motivations or reasons for committing the criminal acts. Mental disorder rarely interferes with the ability to perform the necessary actions to achieve irrationally motivated aims. In cases of self-regulation problems, the defendant does form the mens rea but lacks substantial capacity to conform his conduct to the law.

There are very few contemporary data about the operation of the insanity defense and virtually none about the operation of the mens rea alternative. Montana is the only state for which there is a systematic study of mental disorder claims pre- and post-abolition of the insanity defense.¹⁴⁶ The picture is complicated, but in brief, the number of cases, the types of defendants, and the types of crimes did not change.¹⁴⁷ There were two major effects, however. Under the mens rea alternative, more defendants were convicted, and the number of defendants found incompetent to stand trial markedly increased.¹⁴⁸ Conviction is unjust in any case in which the defendant should have been acquitted by reason of insanity. The increase in convictions in Montana demonstrates that abolition of the insanity defense does, in fact, expose severely mentally ill offenders to unjust punishment. Moreover, the rise in the number of defendants found incompetent to stand trial who would previously have been found competent and acquitted suggests that an incompetency finding is being used as a tool for diversion in cases involving less serious charges that likely would have led to stipulated insanity acquittals under the pre-abolition statute. This is also objectionable.

Although Kahler seemed to have had the mens rea required for the charged offense of capital murder, his expert evidence established that he suffered from a number of mental disorders, including a major mental disorder, severe depression.¹⁴⁹ As a result, the expert opined, Kahler's perception and judgment were so distorted that he may have become dissociated from reality at the time of the crime.¹⁵⁰ The expert also testified that Kahler could not refrain from his conduct.¹⁵¹ Because an insanity defense was not available and

145. *See supra* Part II.B.

146. Lisa A. Callahan et al., *The Hidden Effects of Montana's "Abolition" of the Insanity Defense*, 66 PSYCHIATRIC Q. 103 (1995).

147. *Id.* at 107–12.

148. *Id.* at 115–16.

149. *Kahler v. Kansas*, 140 S. Ct. 1021, 1027 (2020).

150. *State v. Kahler*, 410 P.3d 105, 114 (Kan. 2018).

151. *Id.*

Kahler's conduct met the criteria for capital murder, his conviction for the most serious crime in the criminal law was improperly a foregone conclusion. Justice Breyer's dissent in *Kahler* fully understood the failure of the mens rea alternative as a result of how mental disorder affects mens rea, but to no avail.¹⁵²

We have seen that a delusional belief is objectively unreasonable, so a delusional defendant charged with a potential negligence mens rea, such as homicide, might be convicted on a negligence theory. But convicting the severely disordered defendant of a crime based on a negligence standard is fundamentally unjust, as even Mr. Justice Holmes recognized in his rightly famous essays on the common law.¹⁵³ The defendant's unreasonable mistake was not an ordinary mistake caused by inattention, carelessness, or the like. Defendants are responsible for the latter because we believe that they had the capacity to behave more reasonably by being more careful or attentive. In contrast, the hypothetical defendant's delusional "mistake" was the product of a disordered mind, and thus he had no insight and no ability to recognize the gross distortion of reality. He was a victim of his disorder, not someone who deserves blame and punishment as a careless perpetrator of involuntary manslaughter. He does not deserve any blame and punishment, and only the defense of legal insanity could achieve this appropriate result. Paradoxically, such a defendant's potential future dangerousness if he remains deluded would be better addressed by an insanity acquittal and indefinite involuntary commitment, a practice this Court has approved,¹⁵⁴ than by the comparatively short, determinate sentences for involuntary manslaughter.

Thus, the mens rea alternative is not an acceptable replacement or substitute for the insanity defense. Only in the exceedingly rare case in which mental disorder negates all mens rea would the equivalent justice of a full acquittal be achieved, albeit for a different reason. But again, this is the rarest of cases. Most legally insane offenders form the mens rea required by the definition of the charged offense and only the defense of legal insanity can respond justly to their lack of blameworthiness. Contrary to the claim of the

152. The facts in *Kahler* were not nearly as favorable to an insanity defense as those in an earlier case in which the Supreme Court denied certiorari, *Delling v. Idaho*, 568 U.S. 1038 (2012). *Delling* was clearly delusional and formulated his plan to kill intentionally based on those delusions. *State v. Delling*, 267 P.3d 709, 720 (Idaho 2011). Although *Delling* possessed the mens rea for murder and was found guilty, the trial judge opined that *Delling* did not know right from wrong. *Id.* It was the perfect case to illustrate the failure of the mens rea alternative, but over a written dissent from Justice Breyer, the Court ducked the issue. *Delling*, 568 U.S. at 1039 (Breyer, J., dissenting from denial of certiorari). Justice Breyer used the same argument he adopted in *Kahler*.

153. HOLMES, *supra* note 1, at 50–51.

154. *Jones v. United States*, 463 U.S. 354 (1983).

majority, the Kansas rule does not genuinely address moral culpability.¹⁵⁵ Many people with full mens rea are not morally culpable and are not blameworthy even though they are cognitively capable of forming mens rea. Considering mental disorder for purposes of mitigation at sentencing and the possibility of hospitalization rather than imprisonment after convictions are insufficient, and ersatz moral responses are no substitute for the affirmative defense of legal insanity.

On moral grounds, it is unfair to blame and punish a defendant who deserves no blame and punishment, even if the offender's sentence is reduced. Blaming and punishing in such cases is unjust, full stop. Sentencing judges might also use mental disorder as an aggravating consideration because it might suggest that the defendant is especially dangerous. Sentences of severely mentally ill offenders might be enhanced. Again, injustice would result, and public safety would not be protected as well as by an indeterminate, post-acquittal commitment. Third, unless a sentencing judge is required by law to consider mental disorder at sentencing, whether the judge does so will be entirely discretionary. Again, this is a potential source of substantial injustice. Finally, in some jurisdictions the sentencing judge in cases of serious crime has little if any sentencing discretion, so even the most sensitive, sympathetic judge would be prevented from imposing a just sentence. In short, only a required insanity defense would ensure that arguably blameless mentally disordered offenders have an opportunity to establish that state blame and punishment are not justified.

D. Powell Postscript

In *Powell v. Texas*, Justice Marshall's plurality opinion rejected a claim that the Constitution requires an affirmative defense if a defendant's crime was compelled by a disease.¹⁵⁶ In a part of the decision that the *Kahler* majority cites with approval, Marshall claimed that it would be unwise to constitutionalize such a uniform defense because this should be the province of individual jurisdictions.¹⁵⁷ There was no need in *Kahler* to constitutionalize any particular form of the insanity defense. States could have been left free to adopt any plausible rule. One of the defects of the *Kahler* dissent, in my opinion, is that it suggested that some form of "moral understanding" test was required, thus opening itself to the charge of adopting a one-size-fits-all test

155. See *Kahler v. Kansas*, 140 S. Ct. 1021, 1031 (2020).

156. 392 U.S. 514, 536–37 (1968).

157. *Id.* at 536.

that *Powell* abjured.¹⁵⁸ Some affirmative defense test is necessary, however, and all are interpretively capacious. Consequently, even though I am a staunch opponent of control tests for legal insanity and generally, even a sole control test would be preferable to no test at all. It would be an unwise rule, and I hope legislatures and state courts will not impose control tests in any context in which they wish to broaden excusing conditions.

III. THE EXTERNAL CHALLENGES TO RESPONSIBILITY FROM BEHAVIORAL NEUROSCIENCE & GENETICS

As millennia of philosophizing attest, there are challenging questions about the existence, source, and content of meaning, morals, and purpose in human life, but present and foreseeable neuroscience will neither obliterate nor resolve them. Neuroscience, for all its astonishing recent discoveries, raises no new challenges in these domains. It poses no unique threat to our life, hopes, or to our ability to decide how to live and how to live together, including the use of law to guide us. The supposed challenges were best summed up by an editorial warning in *The Economist*: “Genetics may yet threaten privacy, kill autonomy, make society homogeneous and gut the concept of human nature. But neuroscience could do all of these things first.”¹⁵⁹

The primary quarries of those who think that neuroscience poses a challenge to meaning, morals, and purpose are the related concepts of responsibility and desert, especially as they play a role in the criminal law. After all, responsibility and desert are intrinsic features of present moral and criminal legal concepts, practices, and institutions, including the imposition of punishment. Most of those who challenge responsibility and desert think that these concepts are philosophically questionable and lead to primitive, prescientific practices, such as overly harsh punishments. Responsibility and desert are also intrinsic to civil law, but virtually none of these challengers consider how their views would affect desert theories in contracts, torts, and property law, for example. Although critics have a duty to embed their criticisms of criminal law in a wider understanding of the implications of the criticisms, this Article will nonetheless engage with the dominant critique by

158. *Kahler*, 140 S. Ct. at 1040 (Breyer, J., dissenting). I’m not sure that it would have made a difference if the dissent had conceded this point. The majority was able to avoid the issue by wrongly claiming that Kansas had indeed adopted a form of legal insanity by implementing the mens rea alternative as it did. *Id.* at 1030 (majority opinion).

159. *Open Your Mind*, *ECONOMIST* (May 23, 2002), <https://www.economist.com/science-and-technology/2002/05/23/open-your-mind> [<https://perma.cc/Q7SW-Y92Z>].

limiting itself to the potential effect of neuroscience on the criminal law and moral responsibility more generally.

As is well known, the primary challenges neuroscience allegedly presents to responsibility, desert, and retributive justifications of punishment are the threat from determinism and the specter of the person as simply a “victim of neuronal circumstances” (VNC)¹⁶⁰ or just a “pack of neurons” (PON).¹⁶¹ Allegedly, no one is responsible for any of his behavior, and no one deserves a proportionate response to his behavior either because determinism is true and inconsistent with responsibility or because mental states are epiphenomenal, and we are therefore not the sort of creatures that can be guided by reason.¹⁶² Such criticism based on the claim that no one is genuinely responsible for wrongdoing assumes that a partial or wholly responsibility-based system is incoherent because it rests on a fundamental mistake. The determinist and VNC/PON challenges are thus external critiques of criminal justice. As proponents of these challenges fully recognize, they provide no basis whatsoever for internal reform of a responsibility-based criminal justice system. Determinism is not selective or partial.¹⁶³ If it grounds a moral or legal practice, it applies to all who come within the practice, and it cannot make the distinctions concerning guilt and desert that are at the heart of criminal justice. The same is true of the VNC/PON challenge. It denies the possibility of responsibility, applies to all, and would entail abandoning the moral responsibility distinctions our system now makes using mens rea and affirmative defense doctrines.

As is apparent, the consequences of accepting these critiques would be nothing short of radical and completely unmoored from standard views of responsibility. This is no reason not to adopt such changes if they are justified, but it seems clear that the burden of proof should be placed on the proponents of radical change. They seek to abandon a system that has evolved for centuries; that is in accord with common sense and with moral, political, and legal theories that are widely endorsed; and that seems to work, albeit imperfectly.¹⁶⁴ These challenges cannot borrow internal reformist changes because they are total critiques. If no one is really morally responsible, and consequently, no one deserves any blame or punishment, then only a radically

160. Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANSACTIONS 1775, 1781 (2004).

161. FRANCIS CRICK, *THE ASTONISHING HYPOTHESIS: THE SCIENTIFIC SEARCH FOR THE SOUL* 3 (1994).

162. JOHN MARTIN FISCHER & MARK RAVIZZA, *RESPONSIBILITY AND CONTROL: A THEORY OF MORAL RESPONSIBILITY* 14 (2000).

163. *See id.*

164. For a defense of this stance, see ALASDAIR MACINTYRE, *AFTER VIRTUE* 165–80 (3d ed. 2007).

new system of social control is justified. What reason would there be to replace the time-tested system based on an unresolvable metaphysical argument (incompatibilist) or an unproven scientific (VCN/PON) claim?

Proponents of the challenges often claim that responsibility and desert theorists should bear the burden of persuasion because they are justifying harsh treatment. But this counterclaim begs the question. It assumes that all punishment is unjustifiably harsh because no one deserves any punishment at all, but that is precisely what desert theorists deny. More important, until the brave new world the radical challengers propose is fully described, it is not clear that it is either workable or that it will not be even more inconsistent with human flourishing than the current system based on responsibility and desert.

A. *The Determinist Challenge*

In one form or another, the challenge from determinism to “free will” and responsibility has been mounted for millennia. Neuroscience poses no new challenge in this respect. No science can prove the truth of determinism, and the answers to the determinist threat are the same. People care about the issue because it allegedly underwrites conceptions of responsibility, agency, and dignity that are crucial to our image of ourselves and to our political and legal practices and institutions. The alleged incompatibility of determinism and free will and responsibility is therefore a foundational metaphysical and moral issue. Determinism is not a continuum concept that applies to various individuals in various degrees. There is no partial or selective determinism. If the universe is deterministic or something quite like it, responsibility is possible, or it is not. If human beings are fully subject to the causal laws of the universe, as a thoroughly physicalist, naturalist worldview holds, then many philosophers claim that “ultimate” responsibility is impossible.¹⁶⁵ On the other hand, plausible “compatibilist” theories suggest that for many different reasons, responsibility is possible in a deterministic universe even if no human being has the godlike, contra-causal, ultimate freedom that incompatibilists require.¹⁶⁶ Compatibilists hold that human beings possess whatever degree of freedom or control is necessary for responsibility.

165. See, e.g., DERK PEREBOOM, *LIVING WITHOUT FREE WILL* 46–48 (2001); Galen Strawson, *Consciousness, Free Will, and the Unimportance of Determinism*, 32 *INQUIRY* 3, 15 (1989).

166. See MICHAEL S. MOORE, *MECHANICAL CHOICES: THE RESPONSIBILITY OF THE HUMAN MACHINE* (2020); KADRI VIHVELIN, *CAUSES, LAWS, AND FREE WILL: WHY DETERMINISM DOESN'T MATTER* (2013); WALLACE, *supra* note 139; CHRISTIAN LIST, *WHY FREE WILL IS REAL* (2019).

Compatibilism is the dominant view among philosophers of responsibility, and it most accords with common sense. Much about who we are and what we do is not a product of our rational action, including our genetic endowment, early environment, and the opportunities that present themselves to us. Luck clearly plays an immense role in human life.¹⁶⁷ Nonetheless, the compatibilist claims that we retain sufficient capacity to be guided by reason and to choose otherwise when we act or omit.¹⁶⁸ This is our ordinary view of ourselves and our agency. When any theoretical notion contradicts common sense, the burden of persuasion to refute common sense must be very high, and no metaphysics that denies the possibility of responsibility exceeds that threshold.

There seems no resolution to the incompatibilism/compatibilism debate in sight, but our moral and legal practices do not treat everyone or no one as responsible. Determinism cannot be guiding our practices. If one wants to excuse people because they are genetically and neurally determined or determined for any other reason, one is committed to negating the possibility of responsibility for everyone.

Our criminal responsibility criteria and practices have nothing to do with determinism or with the necessity of having so-called “free will.”¹⁶⁹ The metaphysical libertarian capacity to cause one’s own behavior uncaused by anything other than oneself, the strongest conception of free will, is neither a criterion for any criminal law doctrine nor foundational for criminal responsibility. Criminal responsibility involves evaluation of intentional, conscious, and potentially rational human action. And few participants in the debate about determinism and free will and responsibility argue that we are not conscious, intentional, and potentially rational creatures when we act. The truth of determinism does not entail that actions and nonactions are indistinguishable and that there is no distinction between rational and non-rational actions or compelled and uncompelled actions. Our current responsibility concepts and practices use criteria consistent with and independent of the truth of determinism.

In short, the hard determinist incompatibilists are proposing a radical revision of practices and institutions that have evolved for centuries on the basis of an unresolvable metaphysical claim and in the face of common sense. The case simply is not proven. But, in my view, because this viewpoint makes

167. ROBERT H. FRANK, *SUCCESS AND LUCK: GOOD FORTUNE AND THE MYTH OF MERITOCRACY* 8 (2016).

168. J.G. Moore, *Criminal Responsibility and Causal Determinism*, 9 WASH. U. JURIS. REV. 43, 48–49 (2016).

169. Stephen J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology*, 25 BEHAV. SCIS. & L. 203, 220 (2007).

no obvious internal mistakes, it must be taken seriously, at least in so far as proponents offer a different vision of social ordering rather than just a pure metaphysical argument. Rather than attempt to resolve the unresolvable, it assumes that a compatibilist also makes no internal mistakes and turns to a comparative analysis of the world as we know it and the world that hard incompatibilists are offering us. But first let us briefly consider the most radical claim that neuroscience arguably presents: VNC/PON.

B. The Truly Radical Challenge to Responsibility and Law

This section addresses the hyper-reductive claim that the new sciences, and especially neuroscience, will cause a paradigm shift in our view of ourselves as agents who can direct our own lives and can be responsible by demonstrating that we are “merely victims of neuronal circumstances” or a “pack of neurons” (or some similar claim that denies human agency).¹⁷⁰ This claim holds that we are not the kinds of intentional creatures we think we are. If our mental states play no role in our behavior and are simply epiphenomenal, then traditional notions of agency and responsibility based on mental states and on actions guided by mental states would be imperiled. More broadly, neurons, or even a big pack of them like the connectome, do not have meaning, morals, and purpose. They are just biophysical mechanisms. They are not agents. Meaning, morals, and purpose are products of people, not mechanisms. But is the rich explanatory apparatus of agency and intentionality simply a post hoc rationalization that the brains of hapless homo sapiens construct to explain what their brains have already done? Will the criminal justice system as we know it wither away as an outmoded relic of a prescientific and cruel age? If so, the criminal law is not the only area of law in peril. What will be the fate of contracts, for example, when a biological machine that was formerly called a person claims that it should not be bound because it did not make a contract? The contract is also simply the outcome of various “neuronal circumstances.”

Before continuing, we must understand three things. The compatibilist metaphysics discussed above does not save agency, responsibility, morals, meaning, and purpose if the radical claim is true. If determinism is true, two states of the world concerning agency are possible: agency and all that it entails exists or it does not. Compatibilism assumes that agency is true because it holds that agents can be responsible in a determinist universe. It thus essentially begs the question against the radical claim. If the radical claim is true, then compatibilism is false because no responsibility is possible

170. Greene & Cohen, *supra* note 160, at 1781; CRICK, *supra* note 161, at 3.

if we are not agents. It is an incoherent notion to have genuine responsibility without agency. Second, those forms of hard determinism that accept agency but not responsibility are equally false if the radical claim is true. Finally, this challenge is not new to neuroscience. This type of reductionist speculation long precedes the contemporary neuroscientific era fueled by non-invasive imaging. Neuroscience appears to at last provide the ability to prove the claim by discovering the underlying neural mechanisms that are doing all the work that mental states allegedly do, but the claim is not novel to it. The question remains, however: Is the radical claim true?

Given how little we know about the brain-mind and brain-mind-action connections—we do not know how the brain enables the mind and action¹⁷¹—to claim that we should radically change our conceptions of ourselves and our legal doctrines and practices based on neuroscience is a form of “neuroarrogance.” It flies in the face of common sense and ordinary experience to claim that our mental states play no explanatory role in human behavior, and thus the burden of persuasion is firmly on the proponents of the radical view, who have an enormous hurdle to surmount. There will be attempts to use the new sciences to challenge traditional legal and commonsense concepts, including mens rea, but for conceptual and scientific reasons, there is no reason at present to believe that we are not agents.¹⁷² I should add that similar radical views from psychology, such as “the illusion of conscious will”¹⁷³ and the “automaticity juggernaut,”¹⁷⁴ suffer from the same defects. The burden of persuasion is still firmly on the proponents of the radical view.

Most important, contrary to its proponents’ claims, the radical view entails no positive agenda. If the truth of pure mechanism is a premise in deciding what to do, no particular moral, legal, or political conclusions follow from it.¹⁷⁵ This includes the pure consequentialism that Greene and Cohen incorrectly think follows.¹⁷⁶ The radical view provides no guide as to how one should live or how one should respond to the truth of reductive mechanism. Normativity depends on reason, and thus the radical view is normatively inert. Reasons are mental states. If reasons do not matter, then we have no

171. Adolphs, *supra* note 7, at 175.

172. Stephen J. Morse, *Lost in Translation? An Essay on Law and Neuroscience*, in 13 *LAW & NEUROSCIENCE* 529, 543–54 (Michael Freeman ed., 2011).

173. WEGNER, *supra* note 13.

174. John F. Kihlstrom, *The Automaticity Juggernaut—or, Are We Automatons After All?*, in *ARE WE FREE? PSYCHOLOGY AND FREE WILL* 155–80 (John Baer, James C. Kaufman & Roy F. Baumeister eds., 2008).

175. Mitchell N. Berman, *Punishment and Justification*, 118 *ETHICS* 258, 271 n.34 (2008).

176. Greene & Cohen, *supra* note 160.

reason to adopt any particular morals, politics, or legal rules or to do anything at all.

Even if we still thought that the radical view was correct and standard notions of agency and all that it entails were therefore impossible, we might still believe that the law would not necessarily have to give up the concept of incentives. Such an account would be consistent with “black box” accounts of economic incentives that simply depend on the relation between inputs and outputs without considering the mind as a mediator between the two.¹⁷⁷ For those who believe that a thoroughly naturalized account of human behavior entails complete consequentialism, this conclusion might be welcomed.

On the other hand, this view seems to entail the same internal contradiction just explored. What is the nature of the agent that is discovering the laws governing how incentives shape behavior? Could understanding and providing incentives via social norms and legal rules simply be epiphenomenal interpretations of what the brain has already done? How do we decide which behaviors to reinforce positively or negatively? What role does reason—a property of thoughts and agents, not a property of brains—play in this decision?

As the eminent philosopher of mind and action, Jerry Fodor, reassured us:

[W]e have . . . no decisive reason to doubt that very many commonsense belief/desire explanations are—literally—true.

Which is just as well, because if commonsense intentional psychology really were to collapse, that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species; if we’re that wrong about the mind, then that’s the wrongest we’ve ever been about anything. The collapse of the supernatural, for example, didn’t compare; theism never came close to being as intimately involved in our thought and our practice . . . as belief/desire explanation is. Nothing except, perhaps, our commonsense physics—our intuitive commitment to a world of observer-independent, middle-sized objects—comes as near our cognitive core as intentional explanation does. We’ll be in deep, deep trouble if we have to give it up.

I’m dubious . . . that we *can* give it up; that our intellects are so constituted that doing without it (. . . *really* doing without it; not just philosophical loose talk) is a biologically viable option. But be of good cheer; everything is going to be all right.¹⁷⁸

177. Jack Wiseman, *The Black Box*, 101 ECON. J. 149, 153 (1991).

178. JERRY A. FODOR, *PSYCHOSEMANTICS*, at xii (1987).

Everything is going to be all right. Given what we know and have reason to do, the allegedly disappearing person remains fully visible and necessarily continues to act for good reasons, including the reasons currently to reject the radical view. The law's concepts of mens rea and affirmative defenses and its implicit folk psychology are safe for now and likely to remain that way.

We may be a pack of neurons, but that's not all we are. We are not Pinocchios, and our brains are not Geppettos pulling the strings. And this is a very good thing. Ultimately, I believe that the radical view's vision of the person, of interpersonal relations, and of society bleaches the soul. In the concrete and practical world we live in, we must be guided by our values and a vision of the good life. I do not want to live in the radical's world that is stripped of genuine agency, desert, autonomy, and dignity. I do not want to live in a dystopian society that is stripped of meaning, morals, and purpose. For all its imperfections, the law's vision of the person, agency, and responsibility is more respectful and humane.

IV. CONCLUSION

We are in an era of culpability ferment. Many mens rea reform efforts are meant to make the necessity for genuine culpability more robust by enhancing mens rea, but there are disquieting efforts internal to the criminal law, especially in the context of mental disorder, and there are new, radical challenges both to mens rea and culpability generally arising from new scientific advances. But mental disorder should not be a carveout from responsibility reform. And despite their admirable motivations and attempts to avoid the pitfalls of consequentialism, the radical challenges nonetheless offer a pallid conception of agency and a radical proposal for re-engineering social ordering that is more frightening than our current, imperfect regime. As C.S. Lewis recognized long ago in 1953, a system that treats people as responsible agents is ultimately more humane and respectful.¹⁷⁹

179. See C.S. Lewis, *The Humanitarian Theory of Punishment*, 6 RES JUDICATAE 224, 229–30 (1953).