THE UNEASY ENTENTE BETWEEN INSANITY AND MENS REA BEYOND *CLARK V. ARIZONA*

by

Stephen J. Morse[*] and Morris B. Hoffman[**]

I. Introduction

Ever since the affirmative defense of insanity took its first truly modern breath in 1843 in M'Naghten's Case,[1] its relationship to its cousin, mens rea,[2] has been plagued with confusion. How can one be "insane" yet still have the mental state element required by the definition of a particular crime? Can mens rea alone be a sufficient basis for a sensible theory of criminal responsibility? Are there constitutional limits to a state's power to eliminate or restrict the insanity defense or the mens rea requirement? May states constitutionally preclude defense evidence directly relevant to insanity or lack of mens rea? More generally, what impact should recognized mental disorders have on

---

[*] Ferdinand Wakeman Hubbell Professor of Law and Professor of Psychology and Law in Psychiatry, University of Pennsylvania. I thank Ed Greenlee for his invaluable help. As always, my personal attorney, Jean Avnet Morse, provided sound, sober counsel and moral support.

[**] District Judge, Second Judicial District (Denver), State of Colorado. The views expressed here do not necessarily reflect the views of the Second Judicial District or any of my colleagues on that Court. I want to thank my interns from the University of Colorado School of Law, Edward Hafer and Cameron Munier, for their research and editing help.

[1] Cl. & Fin. 718, 8 Eng. Rep. 718 (H.L. 1843).

[2] "Mens rea" is Latin for "guilty mind." The phrase comes from the English legal precept "*Actus non facit reum nisi mens rea sit*" ["An act is not guilty unless the mind is guilty"], which dates from at least the time of Henry I in the early 1100s, and which was likely based on the writings of St. Augustine. Paul Robinson, Mens Rea, University of Pennsylvania, Scholarship at Penn Law Paper No. 35 (1999), available on line at http://lsr.nellco.org/upenn/wps/papers/35; Francis Bowes Sayre, Mens Rea, 45 HARV. L. REV. 974, 974 (1932). Virtually every civilization that has left a relevant record—including the Babylonians, Jews, Egyptians, Greeks, and Romans—recognized that the law must judge both the act and the intention. See, Radin, Intent, Criminal, 8 ENCYC. SOC. SCI. 126 (1932). This general rule had some strict liability exceptions, such as the quite common ancient rule that a man was strictly liable for the acts of his slaves, and even a strict liability view of homicide. See 2 POLLOCK & MAITLAND, THE HISTORY OF ENGLISH LAW 470-73 (2nd ed., 1968). But these exceptions did not displace the rule that even in ancient times, civilized people cared both about the wrongdoer's acts and his intentions. Indeed, the notion that punishable crimes must be non-accidental appears to be a human universal. See text accompanying notes ** through ** infra. But see Guyora Binder, The New Culpability: Motive, Character, and Emotion in Criminal Law, 6 BUFF. CRIM. L. REV. 1, 15 (2002) ("the supposedly ancient concept of mens rea is really a modern polemical construct").

criminal responsibility? Should evolving ideas about the nature and causes of mental disorders and of human behavior in general require changes in our settled views of blameworthiness?

Last Term, in *Clark v. Arizona*,[3] the Supreme Court had one of its rare opportunities to clarify some of these issues.[4] The particular questions *Clark* presented were whether Arizona's unusually narrow insanity defense test violated the defendant's substantive due process rights and whether an Arizona rule that excluded virtually all expert evidence concerning mental disorder offered for the purpose of negating mens rea violated procedural due process. But alas, the decision clarified little, and may in fact have further muddied the conceptual and practical waters. [5]

To lay the foundation for understanding *Clark*, in Parts II and III we first examine the law of criminal responsibility, including its constitutional dimensions, and then explore the relevance of mental disorder to well-settled principles of criminal responsibility, such as the formation of mens rea and the criteria for the defense of insanity. Part IV considers *Clark* in detail. Part V addresses the constitutionality, and

---

[3] 126 S.Ct. 2709 (2006).

[4] Throughout its history the Court has decided only a handful of cases dealing with broad concepts of criminal responsibility, a reticence that itself reflects the institutional presumption that these questions are largely for the states to decide. See text accompanying notes ** to ** infra. *Clark* was the Court's first case dealing directly with insanity or the general question of mens rea since 1996. See the discussion of Montana v. Egelhoff, 518 U.S. 37 (1996), in the text accompanying notes ** to ** infra. There were mental state cases decided by the Court in the period between *Egelhoff* and *Clark*, but they all dealt either with collateral issues or with a specific type of mens rea. Roper v. Simmons, 543 U.S. 551 (2005) (execution of juvenile who committed capital murder prior to 18 years of age violates Eighth Amendment), discussed in the text accompanying notes ** to ** supra; Atkins v. Virginia, 536 U.S. 304 (2002) (execution of mentally retarded defendant violates Eighth Amendment); Kansas v. Crane, 534 U.S. 407 (2002) (addressing constitutional limits on states' civil commitment powers); Kansas v. Hendricks, 521 U.S. 346 (1997) (same); Holloway v. United States, 526 U.S. 1 (1999) (intent requirement of federal car-jacking statute satisfied by defendant's "conditional intent" to harm driver if necessary).

[5] Accord, Peter Westen, *The Supreme Court's Bout with Insanity:* Clark v. Arizona, 4 OHIO ST. J. CRIM. L.143 (2006). Professor Westen characterizes the issues in *Clark* differently, but he agrees that the opinion failed to clarify them.

wisdom, of abolishing either the requirement of mens rea or the defense of legal insanity. Part VI considers the future of these two doctrines, both of which, we contend, are central to preserving just and sensible principles of blameworthiness. We conclude that contemporary scientific approaches to understanding human behavior do not undermine these principles and our legal practices based on them.

II. Criminal Responsibility and the Constitution

The structure of criminal responsibility is superficially straightforward. Crimes are defined by their "elements," which always include a prohibited act and in most cases a mental state, a mens rea, such as intent. The Constitution's due process clause has been construed to require that the prosecution must prove all the elements defining a criminal offense beyond a reasonable doubt.[6] Even if the state can prove all the elements beyond a reasonable doubt, however, the defendant may avoid criminal liability by establishing an affirmative defense of justification or excuse.[7] The due process clause has been interpreted to permit a state to place the burden of persuasion for affirmative defenses on either the prosecution or the defense.[8]

---

[6]In re Winship, 397 U.S. 358, 364 (1970) ("the Due Process Clause protects the accused against conviction except upon proof beyond a reasonable doubt of every fact necessary to constitute the crime charged."). See also Sandstrom v. Montana, 442 U.S. 510 (1979) (jury instruction shifting burden of persuasion on a mental state element to defendant violates due process).

[7] The border between justification and excuse can be hazy. See, Kent Greenawalt, *The Perplexing Borders of Justification and Excuse*, 84 COLUM. L. REV. 1897 (1984). But the distinction between clear cases may be stated as follows. Conduct that would otherwise be criminal is justified if it is right or at least permissible under the specific circumstances. Self-defense is the classic example. Conduct is excused, even if it is wrongful, if the defendant was not responsible for the conduct. Infancy and legal insanity are classic examples.

[8] Leland v. Oregon , 343 U.S. 790 (1952); (legal insanity); Martin v. Ohio 480 U.S. 228 (1987)(self-defense); Dixon v. U.S., 126 S. Ct. ___ (2006) (duress). It appears that a jurisdiction may place the burden of persuasion on the defendant at any level it wishes, ranging from a preponderance of the evidence to beyond a reasonable doubt. In most cases, however, jurisdictions that shift the burden of persuasion to the

Persistent confusion about the meaning of mens rea, as well as deep and complex constitutional issues, cloud this facially straightforward account of criminal responsibility. First, let us clarify some terminological confusion. In its narrow sense, mens rea refers to the specific mental state or states required by the definition of all crimes other than strict liability crimes.[9] The state must prove these definitional elements beyond a reasonable doubt. Mens rea in this sense is absent, and the defendant must be acquitted of a crime requiring such a mens rea, if the defendant is able to cast a reasonable doubt on whether the mental state was formed in fact.

"Mens rea" is also sometimes used more broadly, however, to refer generally to criminal liability, responsibility, blameworthiness, or culpability. In this sense, it refers to the situation in which the prosecution is able to prove all the definitional elements beyond a reasonable doubt and the defendant fails to establish an affirmative defense.[10] Under these conditions, the defendant is ultimately blameworthy, finally responsible. The defendant can avoid mens rea in this broader sense, defeating any criminal blameworthiness, either by negating any element of the crime charged or by establishing an affirmative defense. In the remainder of this article, we will use the term mens rea in the narrower sense—the specific mental state element required by the definition of the crime.

---

defense require that the defense be proven by a preponderance of the evidence or by the "intermediate," clear and convincing evidence standard. See, e.g., 18 U.S.C. Sec. 17 (placing the burden of proof on defendants in federal trials to prove legal insanity by clear and convincing evidence).

[9] In their influential casebook, Sanford Kadish and Stephen Schulhofer refer to this usage as "special" mens rea. Sanford H. Kadish & Stephen J. Schulhofer, CRIMINAL LAW AND ITS PROCESSES: CASES AND MATERIALS 203 (7th ed., 2001).

[10] Kadish and Schulhofer refer to this usage as "general" mens rea. Id. We use the admittedly vague locution, "establish an affirmative defense," in recognition that the burden of persuasion may be placed on either the prosecution or the defense.

Defining crimes is a quintessential right of states,[11] but many crimes have significant common law, and indeed pre-common law,[12] roots pre-dating the settlements in the Constitution. This raises the central constitutional question about crime: are there constitutional limitations to a state's right to define crime grounded in common law principles codified in the Constitution, either expressly or, more fruitfully, implicitly in the doctrine of substantive due process?[13]

In *Robinson v. California*,[14] the Court held that the Eighth Amendment prohibited punishing a person for being an addict. The Justices disagreed substantially about the rationale for the holding, and there was no single opinion for the Court. The plurality opinion was replete, however, with suggestions that criminalizing and punishing the status of being an addict is unconstitutional because that status is beyond the control of the addict.[15] The *Robinson* Court did not explain how the taking of illegal drugs by an

---

[11] Martin v. Ohio, 480 U.S. 228, 232 (1987); Patterson v. New York, 432 U.S. 197, 201-202 (1977). Indeed, there was little substantive federal criminal law until the early 1900s. See Frank O. Bowman III, *Fear of Law: Thoughts on Fear of Judging and the State of the Federal Sentencing Guidelines,* 44 ST. LOUIS U. L.J. 299, 313-14 (2000).

[12] See text accompanying notes ** to ** infra for a discussion of the biological roots of the criminal law in general and of notions of blameworthiness in particular.

[13] This presupposes another constitutional question, one that is beyond the scope of this article: whether the whole idea of "substantive due process" is legitimate. Compare Mays v. East St. Louis, 123 F.3d 999, 1001 (7th Cir. 1997) (Easterbrook, J.) ("Substantive due process is an oxymoron"), John Harrison, *Substantive Due Process and The Constitutional Text,* 83 Va. L. Rev. 493, 495 (1997) (arguing there is little textual or historical support for the idea of substantive due process), and Robert J. Bork, THE TEMPTING OF AMERICA: THE POLITICAL SEDUCTION OF THE LAW 240 (1990) (same), with Peter J. Rubin, *Square Pegs and Round Holes: Substantive Due Process, Procedural Due Process, and the Bill of Rights,* 103 Colum. L. Rev. 833, 892 (2003) (arguing that Court should read the notion of substantive due process even more broadly); Gregory C. Cook, *Footnote 6: Justice Scalia's Attempt to Impose the Rule of Law on Substantive Due Process, 14* Harv. J. L. & Pub. Pol. 853 (1991) (same).

[14] 370 U.S. 660 (1962).

[15] As Justice Stewart put it, drug addiction "is an illness which can be contracted innocently or involuntarily," and that imprisonment for such a status inflicts cruel and unusual punishment as surely as if the defendant were being imprisoned "for the 'crime' of having a common cold." Id. at 667

addict—which it took great pains to admit could be criminalized by the state—is any more within the addict's control than the status of being an addict.

Six years later the Court stepped back from the precipice of constitutionalizing "loss of control" in *Powell v. Texas*,[16] when it held that the Constitution does not prevent a state from criminalizing public drunkenness, even when the target of the prosecution is suffering from the "disease" of chronic alcoholism and is thus allegedly powerless to control his behavior. Although Justice Marshall's plurality opinion in *Powell* has been read narrowly as simply an unwillingness to extend *Robinson* from the impermissibility of punishing a status to the impermissibility of punishing allegedly uncontrollable acts, the opinion signaled more broadly that the Court considered the doctrines of criminal responsibility firmly within the discretion of the states. It said it was loathe "under the aegis of the Cruel and Unusual Punishment Clause" to become "the ultimate arbiter of the standards of criminal responsibility."[17] Justice Marshall explained:

> We cannot cast aside the centuries-long evolution of the collection of interlocking and overlapping concepts which the common law has utilized to assess the moral accountability of an individual for his antisocial deeds. The doctrines of actus reus, mens rea, insanity, mistake, justification, and duress have historically provided the tools for a constantly shifting adjustment of the tension between the evolving aims of the criminal law and changing religious, moral, philosophical, and medical views of the nature of man. This process of adjustment has always been thought to be the province of the States.[18]

It is difficult to imagine a clearer, more straightforward expression of deference to the states concerning the definitions of crimes and defenses. Yet, beginning in 1975, the Court decided a series of cases that vacillated on the question of the constitutional limits

---

[16] 392 U.S.514 (1968).

[17] *Powell*, n. __ supra, at __. Concurring Justices Black and Harlan agreed.

on a state's freedom to define crimes and defenses and the concomitant power to allocate criminal burdens of proof.

In Mullaney v. Wilbur,[19] the Court considered the constitutionality of Maine's rather unique burden-shifting murder statute. Maine included "malice" as one of the elements of murder, but defined it to exclude killings committed in the heat of passion upon sudden provocation. Even though malice so defined was an element of the crime—an element the prosecution had to prove beyond a reasonable doubt—the Maine statute imposed the burden of proving provocation/passion on the defendant. This shifting of the burden of proof clearly violated *Winship*'s requirement that the state must prove every element of the crime charged beyond a reasonable doubt, and the Court had little trouble striking down the statute on that basis. But the opinion also contained stunning language that seemed to suggest that *Winship* applied to any fact bearing on the defendant's "degree of criminal culpability,"[20] whether or not it was among the elements of the crime.[21] It appeared that the Supreme Court was taking a tentative further step toward constitutionalizing substantive criminal law.

But the imagined turn toward a more intrusive federal role in the substantive criminal law proved just that—imagined. Just two years after the controversial dictum in *Mullaney*, the Court returned to *Powell's* more restrained approach. In *Patterson v. New*

---

[18] Id. at ___.

[19] 421 U.S. 684 (1975).

[20] *Mullaney*, n. __ supra, at __.

[21] This definitional tension between elemental facts and non-elemental facts should have a familiar ring. That same distinction has been the object of the Court's recent and controversial attention in the context of the Sixth Amendment, in the so-called *Apprendi* line of cases culminating in the Court's striking down of the federal sentencing guidelines in United States v. Booker, 543 U.S. 220 (2005).

York,[22] it addressed a New York statute that was functionally equivalent to the Maine statute considered in *Mullaney*. To reduce a killing from murder to manslaughter, New York had adopted the Model Penal Code's "extreme emotional disturbance" test.[23] It was clear that this test was meant to serve the same partially mitigating function as the provocation/passion formula, but it provided a broader test that was allegedly more consistent with modern understanding of behavior and that better tracked culpability. Unlike Maine, however, New York explicitly treated this doctrine as a partial affirmative defense and placed the burden of persuasion on the defendant.

Although "extreme emotional disturbance" is as relevant to the degree of a defendant's culpability as "heat of passion," the Court rejected a broad reading of *Mullaney* that would have required the state to prove beyond a reasonable doubt every fact bearing on culpability, and instead gave determinative weight to New York's decision to treat emotional disturbance as a partial affirmative defense:

> [E]ven if we were to hold that a State must prove sanity to convict once that fact is put in issue [a position the Court rejected], it would not necessarily follow that a State must prove beyond a reasonable doubt every fact, the existence of nonexistence of which it is willing to recognize as an exculpatory or mitigating circumstance affecting the degree of culpability or the severity of the punishment.[24]

The Court strongly re-iterated that states have extensive discretion to define crimes and defenses as they wish and to allocate the burden of persuasion accordingly:

> If the State . . . chooses to recognize a factor that mitigates the degree of criminality or punishment, we think the State may assure itself that the fact has been established with reasonable certainty. To recognize at all a mitigating circumstance does not require the State to prove its

---

[22] 432 U.S. 197 (1977).

[23] Model Penal Code, Sec. 210.3(1)(a)-(b).

[24] Id. at ___.

> nonexistence in each case in which the fact is put in issue, if in its judgment this would be too cumbersome, too expensive, and too inaccurate.[25]

The *Patterson* Court did recognize that there were constitutional limits to the state's ability to redefine at least some elements as affirmative defenses and thus to shift the burden of persuasion, but the examples it gave—such as a legislature declaring a person presumptively guilty of a crime[26]—seemed to place little if any effective limit on the state's discretion.[27]

Patterson did not explicitly overrule *Mullaney*, but it seemed inconsistent with it, and that apparent inconsistency created an industry of commentary.[28] Indeed, the two cases together spawned an enormous debate between the "substantivists" and "proceduralists."[29]

Substantivists believed that *Mullaney* was correctly decided, that *Patterson* threatened to erode *Winship*'s protections, and that states could not, by the mere artifice of re-labeling an element as an affirmative defense, relieve the state of its burden of proof on the definitional elements. For example, imagine a homicide statute that defined "intentional action that caused the death of a human being" as first degree murder and then forced the defendant to prove as affirmative defenses that the defendant did not intend to cause death (reducing the grade to second degree murder), was not aware of a

---

[25] Id. at ___.
[26] Id. at ___.
[27] See Westen, note ** supra, at 161 (because the distribution of elements and defenses is entirely malleable, depending on how states define crimes and defenses, the constitutional rules concerning burden of persuasion are themselves malleable).

[28] See, e.g., Ronald J. Allen, *Structuring Jury Decisionmaking in Criminal Cases: A Unified Constitutional Approach to Evidentiary Devices*, 94 HARVARD L. REV. 321 (1980).

[29] Larry Alexander describes the debate. Larry Alexander, *The Supreme Court, Dr. Jekyll, and the Due Process of Proof*, 1996 SUPREME COURT REV. 191, 193-196.

substantial and unjustifiable risk of death (reducing the grade to manslaughter), and should not have been aware of a substantial and unjustifiable risk of death (acquittal).[30] In such a situation, *Winship*'s protection of liberty would be largely illusory.

In contrast, the proceduralists thought that, as a matter of constitutional law if not of policy, the states have virtually limitless discretion to re-label traditional elemental facts as affirmative defenses. Indeed, that very decision—whether a fact should be an element for which the state has the burden of proof or an affirmative defense that the defendant might be required to prove—is precisely the kind of policy decision the proceduralists believed cases like *Powell* and *Patterson* placed within the state's judgment.

In *Montana v. Egelhoff,*[31] which was the last important general *mens rea* case decided before *Clark*, the Court addressed these issues, but in a context different from re-labeling elements as affirmative defenses. The question answered by *Egelhoff* was whether, and under what conditions, the State may constitutionally bar a defendant from introducing relevant and reliable evidence for the purpose of negating an element. Egelhoff was charged with first degree murder under a Montana homicide statute that defined the crime as requiring purpose or knowledge. Egelhoff claimed that he was so drunk at the time of the killing that he could not have formed a purpose to kill. But under

---

[30] Such a statute would of course raise constitutional eyebrows after Morissette v. United States, 342 U.S. 246 (1952), in which the Court expressed distaste for strict liability applied to core, common law offenses. Nonetheless, Morissette was not constitutionally-based and it is still an open question whether a legislature can choose to make any crime a strict liability crime as long as it makes clear that it intends to do so for good reason. One of us has discussed an even more extreme case of re-labeling elements as non-elements, albeit in the context of A*pprendi v. New Jersey*, 530 U.S. 466 (2000)(holding that the state must charge and prove beyond a reasonable doubt any fact other than prior criminal record that will have the effect of increasing the penalty for a crime beyond the statutory maximum): Could a state define all crimes as having a single element (intending to harm another person) and then change all other elements to sentencing factors? Morris B. Hoffman, *The Case for Jury Sentencing*, 52 DUKE L.J. 951, n. 127 (2003).

a Montana statute that specifically barred the use of evidence of voluntary intoxication to negate mens rea,[32] the evidence of voluntary intoxication was excluded from the jury.

The Supreme Court held that such exclusion did not violate the defendant's due process rights. The justices disagreed about whether the Montana statute was an evidentiary rule or a redefinition of the offense of murder All nine justices agreed, however, that states had substantial discretion to define crimes. As Justice Ginsburg said in her concurrence, citing *Powell*, "States enjoy wide latitude in defining the elements of criminal offenses," particularly when determining "the extent to which moral culpability should be a prerequisite to conviction of a crime."[33] Moreover, all nine justices agreed that a criminal defendant does not have an absolute right to present evidence relevant to his defense.

Justice Scalia began the plurality opinion by noting that there was no indication in historical practice that the use of voluntary intoxication evidence to negate mens rea was so entrenched that it was a "fundamental principle of justice."[34] He continued by arguing that the defendant's right to introduce even "crucial" and reliable evidence was not absolute and could yield to "valid state justifications" that might support exclusion.[35] One

---

[31] 518 U.S. 37 (1996).

[32] MONT. CODE ANN. § 45-2-203 (1995).
[28] *Egelhoff*, supra note ___ at ___, citing Powell v. Texas, supra note ___ at ___ (Black, J. concurring).

[34] An argument could be made that the older authorities, such as Hale, Blackstone and Story, failed adequately to distinguish between the two senses of mens rea described in note ** supra, and that they were referring to the broader sense in rejecting an intoxication defense, But Justice Scalia was correct about the practice in the states. At most, there was a trend in favor of introducing such evidence, albeit almost always with limitations.

[35] Justice Scalia posited several such justifications. First, a state might wish to deter intoxication itself by holding drunken defendants strictly liable for all consequences that occur, even if they lacked mens rea for the aggravated results. Second, the rule operates as a specific deterrent by imprisoning people who cannot control their violent conduct while intoxicated . Third, the rule reflects society's moral perception that a defendant who voluntarily impaired his own faculties should be held strictly accountable for the

might of course take issue with some or all of these justifications, but these kinds of policy decisions are precisely the kinds of decisions arguably left to state legislatures.[36]

The plurality also made short work of Egelhoff's argument that the Montana rule violated *Winship*. Justice Scalia reasoned that Montana's rule did not shift the burden of persuasion at all, but by excluding intoxication evidence simply made it easier for Montana to prove mens rea beyond a reasonable doubt, an effect that any evidentiary rule might produce. Such burden-reducing evidentiary rules are constitutionally permissible unless they themselves violate a fundamental principle of fairness, which, Justice Scalia had already concluded, Montana's rule did not.[37]

This was the rather precarious state of the law until the Court had its chance to speak again in *Clark*. Before we address that opinion, we first briefly consider the concept of mental disorders, and then review in some depth well-settled principles of criminal responsibility and the relevance of mental disorder to those principles.

III. Mental Disorder and Criminal Responsibility

---

consequences. Finally, intoxication evidence may be misleading because there is evidence that culture rather than pharmacology largely explains the connection between crime and intoxication. Jurors who have been led to believe the opposite will be too quick to conclude that intoxicated defendants were "biologically incapable" of forming mens rea. Id. at ___

[36] Justice O'Connor's dissent focused on the "`fundamental principle' that a defendant has a fair opportunity to put forward his defense." *Egelhoff*, supra note ___ at ___ (O'Connor, J., dissenting). She argued that it was fundamentally unfair to prevent the defendant from mounting a full defense with relevant evidence that alone could exonerate him, without close examination of competing state interests, an examination she found wanting. For example, Montana allowed evidence of *involuntary* intoxication to be used to negate mens rea, a rule arguably inconsistent with the claim that intoxication evidence misleads jurors about the formation of mens rea. The dissent also noted that Justice Scalia's policy reasons were not articulated by Montana, either in legislative history or even in the briefs. Montana had instead argued that the "sole purpose" for disallowance was "to keep from the jury's consideration a category of evidence that helps the defendant's case and weakens the government's case." *Id.*

[37] Id. at ____.

What is a "mental disorder" and what is its potential relevance to whether a person suffering from it acted, possessed a requisite mens rea, or qualified for a mitigating or excusing condition such as legal insanity[38]? To answer these questions requires more detail about the nature of recognized mental disorders and the law's criteria for act, mens rea and legal insanity.

A. Mental Disorder

There is no consensus—philosophically or scientifically—on the definition of "mental disorder." That lack of consensus should hardly surprise us. "Disorder" and "order," abnormality and normality, are normative terms applied to behavior that science cannot independently define and that therefore inevitably lead to disagreement. They implicate many difficult empirical and theoretical questions that are relevant to our argument, but are largely beyond its scope. We will later have occasion to examine some of these questions, but for now let us focus on the current scientific and clinical orthodoxy about mental disorder.

Probably the most widely used definition in the United States, and perhaps in the world, is contained in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders.[39] This document admits, however, that, "no definition adequately specifies precise boundaries for the concept of `mental disorder'," [40] and that its definition has been subjected to substantial criticism.[41] The Supreme Court has

---

[38] For purposes of this article, we limit discussion of excuses to legal insanity, but mental disorder might also be relevant to other excuses depending on how the other excuse doctrines are interpreted.

[39] American Psychiatric Association, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS, 4TH EDITION-TEXT REVISION xxx-xxxi (2000)[hereinafter DSM-IV-TR](providing a definition).

[40] Id. at xxx.

repeatedly and opportunistically used this disagreement to justify outcomes as disparate as requiring an intermediate, "clear and convincing evidence" burden of proof in involuntary civil commitment cases,[42] rejecting a constitutionalized control test for criminal responsibility,[43] and upholding the constitutionality of an especially onerous form of involuntary commitment for so-called mentally abnormal sexually violent predators.[44]

Nevertheless, there is substantial agreement that some behavior, especially behavior indicative primarily of a loss of the capacity for rationality, should be considered the product of a biological or psychological defect. For example, most observers agree that a gross and substantial loss of touch with reality qualifies as a severe abnormality. Indeed, there are special legal rules for the differential treatment of people with such abnormalities in most areas of civil and criminal law.[45] Within the latter, for example, in addition to doctrines concerning criminal responsibility, there are doctrines concerning competence to plead guilty, competence to stand trial, competence to be sentenced, and competence to be executed.

There are no doubt many open conceptual and empirical questions about mental disorders, but the law routinely deals with cases involving people who suffer from them. One way to think about these difficult questions about mental state and blameworthiness is to analogize them to electricity. Modern physics has exposed the foundations of the

---

[41] E.g., Jerome C. Wakefield, *Disorder as Harmful Dysfunction: A Conceptual Critique of* DSM-III-R's *Definition of Mental Disorder*, 99 PSYCH. REV. 232 (1993) (critique of same definition used in DSM-IV-TR).

[42] Addington v. Texas, 441 U.S. 418 (1979).

[43] *Powell*, supra note ___.

[44] Kansas v. Hendricks, 521 U.S. 346 (1997).

elemental particles, including the electron, as a kind of mysterious expression of a set of strange and counter-intuitive physical rules; in some ways it seems the more we learn about quantum physics the more our macro-reality seems to be an illusion. Yet those foundational uncertainties hardly disable us from using electricity, or designing circuits.

Healthy skepticism about the relationship between mental disorder and criminal responsibility, especially in marginal cases, is warranted; outright rejection of that relationship is not.

B. The Act Doctrine

An intentional act is crucial to the definition of any conduct we wish to criminalize. Indeed, our very use of the word "conduct" in the last sentence conveys the notion that acts, not mere thoughts or desires, are the domain of crime. Although thoughts and desires can be intentionally produced, and can even be evaluated morally,[46] we do not criminalize these mental states because by themselves they do no harm, and the reach of the criminal law is defined in large measure by the harm principle.[47] On the other hand, harms or threatened harms, all of which result from actual or potential body movements or intentional lack of movement, cannot be attributed to the agent as an agent unless those movements or lack of motion are intentional. Spasms and reflexes, for example, can cause or threaten harm, but in most cases we do not hold the agent responsible for them. The challenge for the law is how to define a "voluntary" or

---

[45] See, Stephen J. Morse, *Crazy Reasons*, 10 J. CONTEMP. LEG. ISSUES 189 (1999).

[46] Merely coveting one's neighbor's wife, for example, violates the Ninth Commandment. Exodus 38:24; Deuteronomy 10:4.

[47] Joel Feinberg, HARM TO OTHERS: THE MORAL LIMITS OF THE CRIMINAL LAW 1-64 (1984).

"intentional" act (or omission in cases in which the agent has a duty to act) in a sensible and useful way.

The criminal law's evolving answer has never been entirely satisfactory. Indeed, some legal philosophers have been forced to define a criminally cognizable "act" by what it is not.[48] Nonetheless, as with the notion of "mental state," the law has managed to develop a useful working definition: an "act" for purposes of the criminal law is an intentional bodily movement (or intentional inaction) performed in a state of reasonably integrated consciousness.[49] "Intentional" in this definition means that the agent's bodily movement is at least potentially rationalizable as having been performed for a reason.

Mental disorder seldom interferes with an agent's ability to act intentionally, even though in infrequent cases it may cloud consciousness. For example, depersonalization and less severe delirium can compromise consciousness without rendering the person's movements non-actions.[50] It is only at the extreme, when consciousness is nearly or completely obliterated, that the law will acquit the defendant because action was lacking Sleepwalking is a classic example.[51] There is a debate about whether apparently environmentally-responsive, goal-directed behavior should be considered action or not,[52]

---

[48] MPC, Sec. 2.01(2).

[49]One must usually infer this from negative criteria. See, eg, MPC, Sec. 2.01(1).

[50] DSM-IV-TR, note *** supra at 135-141 (delirium, point prevalence for which in adults under the age of 55 is .4%); Id at 530-532 (Depersonalization Disorder). Many other disorders can also cause consciousness clouding although such clouding is not a crucial diagnostic criterion.

[51] DSM-IV-TR, supra note *** at 639-644 (Sleepwalking Disorder). See, Fain v. Commonwealth, 78 Ky. 183 (1879).

[52]See, e.g., MICHAEL S. MOORE, ACT AND CRIME 49-52, 135-155, 257-58 (1993) (arguing that cases of compromised consciousness should be treated as non-action); Michael S. Moore, *More on Act and Crime,* 142 U. PA. L. REV. 1749, 1804-20 (1994) (same); Stephen J. Morse, *Culpability and Control*, 142 U. PA. L.

but criminal behavior performed in such states either negates the action requirement of the prima facie case or it satisfies the criteria for the affirmative defense of automatism.[53] Sufficiently clouded consciousness as a result of mental disorder might thus defeat criminal responsibility, but, again, these cases are rare.

C. Mens Rea

As Justice Holmes famously observed, "Even a dog distinguishes between being stumbled over and being kicked."[54] Mens rea is considered the royal road to the assessment of culpability because it indicates the agent's attitude towards the rights and interests of his fellow citizens when the agent harms or threatens them. As Justice Jackson wrote in *Morissette,*

> The contention that an injury can amount to a crime only when inflicted by intention is no provincial or transient notion. It is as universal and persistent in mature systems of law as belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil. A relation between some mental element and punishment for a harmful act is almost as instinctive as the child's familiar exculpatory, "But I didn't mean to"….[55]

Intentional body movements may cause harms, but if the defendant has been acting as carefully as we can expect a morally well-constituted agent to behave under the circumstances, we conclude that the agent has done nothing wrong, civilly or criminally. His behavior has shown all the respect for the rights of others that we demand. Similarly,

---

REV. 1587, 1641-52 (1994) (clouded consciousness should be treated as an affirmative defense); Bernard Williams, *The Actus Reus of Dr. Caligari*, 142 U. PA. L. REV. 1661 (1994) (same).

[53] Wayne R. LaFave, CRIMINAL LAW (4TH ED.) 466-471 (2003).

[54] Oliver Wendell Holmes, Jr., THE COMMON LAW 3 (Dover Ed., 1991).

[55] See Morissette v. U.S., 342 U.S. 246, 250-51 (1952).

an agent who harms another on purpose is more blameworthy and deserves more blame and punishment than an agent who causes precisely the same harm, but does so unaware that his unreasonable conduct was placing the victim at substantial risk of incurring that harm. The disregard of the former for the rights of his victims is total; the disregard of the latter is serious, but less so. Mens rea is crucial to defining fault and justifying punishment.

Over the centuries, the criminal law has developed a very large number of mens rea terms to mark such differences in culpability. Modern codes have attempted to streamline the mens rea vocabulary. The Model Penal Code, for example, has consolidated mens rea to four terms.[56] Nevertheless, the common law process of interpretation continues as courts struggle to find means more accurately to evaluate blameworthiness.

With rare exception, mens rea terms, such as intent or knowledge, have their ordinary, non-technical meanings. For example, intent does not require any level of commitment to the agent's goal or any degree of moral evaluation.[57] Intent simply means that the agent did something on purpose. Even where a mens rea term is legal jargon,

---

[56] Those four levels of culpability, in order of decreasing blameworthiness, are: purpose, knowledge, recklessness, and negligence. MODEL PENAL CODE § 2.02(2) (1985). Several scholars have argued that these apparently distinct levels of culpability are in fact indistinguishable from one another, at least at their margins. See, e.g., Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Responsibility,* 88 CAL. L. REV. 931 (2000*)*; Douglas N. Hasak & Craig A. Callender, *Willful Ignorance, Knowledge and the "Equal Culpability" Theory: A Study of the Deeper Significance of the Principle of Legality,* 1994 WIS. L. REV. 29 (1994*)*; Kimberley K. Ferzan, *Opaque Recklessness,* 91 J. CRIM. L. & CRIMINOLOGY 597 (2001*)*; Morris B. Hoffman, Booker, *Pragmatism, and the Moral Jury*, 13 GEO. MASON L. REV. 445, 473-74 (2005).

[57]See, Sanford Kadish & Stephen Schulholfer, CRIMINAL LAW AND ITS PROCESSES (7TH ED.) 203 (2001). There are some exceptions. Premeditation, which is used in some jurisdictions as a criterion for the most serious form of intentional murder, under some interpretations may include the requirement of a particularly rational or even morally evaluative process of preconsideration before committing intentional homicide. Negligence standards employing a "reasonableness" test of course require normative judgment about whether behavior is reasonable.

such as "recklessness," its criteria—conscious awareness of a substantial and unjustifiable risk—have traditional meanings. To determine if mental disorder negates mens rea, one must simply ask, using straightforward common sense to provide an answer, if the defendant's disordered mental state actually indicates that mens rea was not formed on the occasion.

Doctrines that permit defendants to present mental disorder evidence to negate mens rea are often misleadingly termed "diminished capacity," suggesting that they are some kind of mitigation or partial excuse. This is incorrect. These doctrines refer to a straightforward denial of a requisite element, akin to a denial that one acted or to a mistake of fact.[58] 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 introduce mental disorder evidence to negate mens rea because courts wrongly believe the defendant is requesting the court to create a "mini" insanity defense.[59]

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 reasons for action, but it virtually never prevents a defendant from meeting the law's criteria for intention, knowledge, conscious awareness, and other mens rea terms.[60]

---

[58] Stephen J. Morse, *Undiminished Confusion in Diminished Capacity*, 75 J. CRIM. L. & CRIM.. 1 (1984).

[59] E.g., Bethea v. U.S., 365 A.2d 64 (D.C. 1976). The Arizona case the Supreme Court considers in *Clark* makes precisely this error. State v. Mott, 186 Ariz. 536, 540-41, 931 P.2d 1046 1050-51 (1997). Indeed, as we shall see, the *Clark* majority itself blurs the two doctrines. See text accompanying notes ** to ** infra.

[60] See Morse, supra note **, at ___.

Likewise, mental disorder seldom disables a defendant from having the *capacity* to form a mens rea. Modern inquiries into state 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 intended a result 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, but failure to have a mens rea, or having a mental state inconsistent with the requisite mens rea, does not mean 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. Thus evidence about the defendant's capacity to form a mens rea is logically and factually relevant to whether it was actually formed. But asking about a defendant's capacity to form a mental state never provides better information than inquiring directly whether the requisite mens rea was formed in fact, which is the ultimate legal question.

What is more, resolving questions about capacity requires a counterfactual inquiry that we lack the clinical and scientific resources to answer. Most of the time when an expert testifies that a defendant lacked the capacity to form a mens rea, that opinion has no clinical or scientific basis. It was precisely these types of difficulties that led California to bar testimony about the defendant's capacity to form mens rea although it did permit testimony about whether the mental state was formed in fact.[61]

Consider the Model Penal Code's example of a hallucinating killer who strangles a victim believing that he is squeezing a lemon.[62] Believing that you are squeezing a lemon is flatly inconsistent with believing that you are squeezing someone's throat for

---

[61] Cal. Penal Code, Sec. 28

the purpose of killing the victim. But what does it mean to ask whether such a person had the capacity to form the intent to kill? How do we interpret this question?

One possibility is that we are asking whether the person could have formed an intent to kill if he believed that he was squeezing a person's throat, whether, that is, he had the general capacity to form an intent to kill a person. This requires a speculative answer, but the answer is probably "yes," because most people have the general capacity to form all species of intents, including an intent to kill.

The second possible interpretation is that we are asking whether at the precise moment that a person believes he is squeezing a lemon, he is also capable of forming an intent to kill the object he is squeezing. The answer to this second question is "no," because one cannot at precisely the same moment form two different and inconsistent mental states. This is tautologically and trivially true. But this shows that we are not really considering the defendant's capacities at all. The reason we know that he "lacked the capacity" to form the intent to kill at that precise moment is because he in fact formed another mental state. Our knowledge of this very temporally specific form of "incapacity" is doing no explanatory work and is a conclusion following from our ex post knowledge of what mental state was in fact formed.

So far, the discussion of capacity has assumed that we incontrovertibly know what mental state the defendant in fact formed. In the real world of criminal law, however, mental state must be proved by the prosecution, and is often countered by defense evidence. To continue the lemon example, imagine that there was good evidence to suggest that the defendant sporadically hallucinated that non-lemons were in fact

---

[62] MODEL PENAL CODE § 4.01, commentary at 156 (1955).. This example is silly because mental disorder does not work this way; it does not produce such beliefs. But the virtue of the example is that it is clear.

lemons, but it is not clear whether he was suffering from that hallucination when he strangled his victim. He claims that he was, but there was also evidence of bad blood between him and the victim.[63] Perhaps he is lying, or he is so distraught about committing homicide that he has unwittingly and incorrectly convinced himself that he was hallucinating.

Only when we are confident that he almost always hallucinates that people are lemons can we conclude with some confidence that he probably lacked the capacity to believe that the victim was not the small yellow citrus fruit. And again, our conclusion about capacity would be based on what mental states had been formed in fact rather than based on our evidence about the person's capacities. The mental states of people with mental disorder, like those of people without disorder, are often variable, however, and speculation about capacity on a particular occasion will be of little help.

For all of these reasons, evidence of mental disorder allegedly relevant to a person's capacity to form a mental state will almost never be helpful about whether that state was formed in fact.

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. When he shot and killed Peel's secretary, Drummond, believing the secretary was Peel, he surely intended to kill. Likewise, Andrea Yates believed that unless she killed her children, they would

[63] This was almost precisely the situation in *Clark*. There was no doubt Clark was suffering from paranoid schizophrenia, but considerable doubt about whether his delusions really interfered with his ability to intend to kill a policeman. Just days before he had bragged to his friends about wanting to kill policemen. See text accompanying notes ** to ** infra.

become corrupt and would be tormented by Satan for eternity. She therefore decided to kill her children. She knew they were human beings and that human beings are killed by drowning. When she drowned the five children in the bathtub, she surely intended to kill them. Likewise, a person suffering from hallucinations who hears God's voice command him to kill surely forms the intention to kill when he kills in response to the hallucinated command.

Admittedly, on rare occasions, psychotic mentation is genuinely inconsistent with the formation of mens rea. In a well-known California case,[64] the defendant, Wetmore, was caught in the victim's apartment under conditions suggesting that he intended to steal the victim's property. Charged with burglary, Wetmore claimed that he delusionally believed that the apartment and the property belonged to him. If he was telling the truth, he did not intent to enter the apartment of *another* or to commit the felony of larcency, 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. The person with mental disorder who is unaware of a risk that a reasonable person should be aware of is by definition unreasonable. 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 situation*" before making a finding of negligence,[65] 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 makes clear that it wishes to

[64] People v. Wetmore, 22 Cal. 3d 318 (1978).

avoid complete subjectivization of the reasonable person standard.[66] To assess reasonableness from the standpoint of the "reasonable irrational" person would deprive the negligence standard of all objectivity.[67] Consider the lemon example again. The killer is at least guilty of negligent homicide because a reasonable person should have been aware of a substantial and unjustifiable risk that the object being squeezed was a person's throat and not a lemon. We defer discussion of whether a conviction and punishment for negligent homicide in such a case would be fair or even make sense until a later section of the article.[68]

In other cases, mental disorder may not necessarily be inconsistent with formation of mens rea, but 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 simply broke in to keep warm.

Again, the crucial issue is to determine what the defendant's actual mental state was and to compare that mental state to the mental state required by the crime charged. Of course, the problem lurking among these theoretical niceties is that it is very difficult,

---

[65] MPC, Sec. 2.02(2)(d) (emphasis supplied)

[66] MODEL PENAL CODE AND COMMENTARIES, PART I 242 (1985).

[67] Failure to recognize this point and the incorrect belief that the MPC adopts nearly complete subjectivization for negligence are major analytic reasons that the primary contemporary proposal to abolish the insanity defense fails. See, Christopher Slobogin, *An End to Insanity: Recasting the Role of Mental Disability in Criminal Cases* 86 VIRGINIA L. REV. 1199, 1239 (2000), which is reprinted in expanded form, but with unchanged analysis, in, Christopher Slobogin, MINDING JUSTICE: LAWS THAT DEPRIVE PEOPLE WITH MENTAL DISABILITY OF LIFE AND LIBERTY 52-53 (2006). Professor Slobogin's "integrationist" alternative to the insanity defense is discussed in the text accompanying notes **--** infra.

[68] Part ___ infra.

or it is at least claimed by some commentators and all defense attorneys to be very difficult, to determine a defendant's mental state at the time of the crime. Memories fade or are wittingly or unwittingly shaped by what happens afterwards. Defendants have powerful incentives to lie. But these difficulties arise in all retrospective mental state evaluations, and not just in cases involving mental disorder. As we discuss later in this article, one of the primary skills we have evolved over the last 100,000 years is the ability to sense the intentions of other humans.[69] Like Holmes' dog, jurors know the difference between a kick and a stumble.

D. Legal Insanity

The tests for legal insanity may be placed in two broad categories—cognitive and control (also called "volitional.") There is some doctrinal variation among the jurisdictions, but the essential questions are similar.

Cognitive tests ask whether at the time of the crime, the defendant was disabled by any mental disorder from knowing, or appreciating or understanding the nature of his conduct or that it was morally or legally wrong. Control tests ask whether, as a result of mental disorder, the defendant was unable at the time of the crime to control his conduct or to conform his conduct to the requirements of the law. [70]

Prior to the 1980s, all American states and all federal circuits had an insanity defense test. The most common was M'Naghten or some variant of its cognitive inquiry.

---

[69] See text accompanying note ** to ** infra.

[70] Sometimes the tests for insanity are phrased instead as tests for "sanity," in which case, of course, the disjunctive becomes conjunctive. Thus, the cognitive test for sanity is whether a defendant understands the nature of his actions *and* knows they are wrong. (These are the two traditional M'Naghten prongs, one of which was dispensed with by Arizona, see text accompanying note ** to ** infra.)

Beginning in the last part of the 19th Century, a minority of jurisdictions added a control test, but no jurisdiction had a control test alone. The Model Penal Code promulgated an influential test that had both a cognitive and a control prong, and that used less all-or-none, bright-line language than standard insanity tests.[71] By the time John W. Hinckley was acquitted by reason of insanity in 1982 for attempting to assassinate President Reagan and others, most states that had reformed their insanity defense followed the Model Penal Code, as had all but one federal circuit.

Hinckley was acquitted under the Model Penal Code test, which spurred an immense federal and state legislative backlash against the insanity defense in general and control tests in particular. In the wake of Hinckley, many states reverted to a pure cognitive test**,** and Congress adopted such a test, providing for the first time a uniform test for legal insanity in federal trials.[72] In addition, Congress and many state legislatures placed the burden of persuasion on the defendant.[73] Five states abolished the insanity defense altogether, although one State supreme court found the abolition unconstitutional.[74] The Supreme Court of the United States has never ruled, however, whether the insanity defense is or is not constitutionally required.

None of the cognitive or control tests for legal insanity is self-defining. All require substantial interpretation. To begin, how should the threshold requirement of a

---

[71] MODEL PENAL CODE, Sec. 4.01(1): A person is not responsible for her criminal conduct if at the time of the crime, "as a result of mental disorder or defect the defendant lacks substantial capacity either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of the law."

[72] 18 U.S.C.. § 17(a). "It is an affirmative defense…that, a the time of the commission of the acts constituting the offense, the defendant, as a result of severe mental disease or defect, was unable to appreciate the nature and quality or the wrongfulness of his acts."

[73] Id.; e.g., ARIZ. REV. STAT. ANN., Sec. 13-502(A) (West 2001).

mental disease or disorder or defect be defined? The Supreme Court has made clear that the definition of mental disorder as a legal criterion for insanity is a legal question and within the discretion of the states. Thus, states are free to limit legal insanity to cases involving severe mental disorder or to define mental disorder in a non-traditional way, as long as the definition is minimally rational. They need not be bound by the definitions used by the mental health science disciplines, such as psychiatry and psychology.[75]

Moreover, there is substantial leeway in whether the tests should be read narrowly or broadly. Consider whether Andrea Yates knew what she was doing when she drowned her five children to save them from Satan's eternal torments. She did know that they were children, that drowning would kill them, and that killing them violated the law and the dominant view of morality, but she also delusionally believed that homicide was necessary to save the children from a fate far worse than premature death. In a narrow sense, she did know what she was doing and that it was morally and legally wrong. In a broader sense, however, and assuming we believe her and her expert psychiatrists, she did not know what she was doing because she was genuinely motivated by an utterly mistaken belief that was the product of some profound abnormality, and she plausibly would have been justified if the belief were true. Using words like "appreciate" or "understand," rather than "know," in a cognitive test will not help. They may signal that the test is to be interpreted more broadly, but how much more broadly and in what way?[76]

---

[74] See notes ** to **infra.

[75] *Hendricks*, supra note ___, at ___.

[76] See, Douglas Mossman, United States v. Lyons: *Toward a New Conception of Legal Insanity*, 16 BULLETIN OF AM. ACAD. OF PSYCHIATRY AND LAW 49, 54-57 (1988). Dr. Mossman provides a thoughtful definition of "appreciate" in the context of the insanity defense, but the term does not entail his usage and the criteria he employs could by interpretation be made part of the criteria for "knowledge" or "understanding." Moreover, a court could interpret the term "appreciate" in the same general way as

Control tests are even more difficult to apply. We all have a common sense understanding of when someone's perceptions or beliefs are irrational, but what does it mean to say that a person cannot control himself? If the person is having difficulty conforming to law because he is irrational, then irrationality is doing the potentially excusing work and there is no need for a separate control test.[77] For example, if Ms. Yates believed that she had an absolute moral obligation to "save" the children by drowning them, it would indeed be difficult for her to obey the law, but the problem is her irrational belief, not a failure of self-control. In fact, lack of rationality explains most cases of failure to conform to law among people with severe mental disorder.

Control tests are often characterized as volitional, as if mental disorder has somehow disabled the agent's volitional capacities, but this locution is confusing and should be abandoned. There is no consensual definition of the volition or will in any of the relevant disciplines, such as psychology or philosophy, and this formulation quickly dissolves into the intractable issues of how the brain enables the mind and how intentions are related to actions.

Perhaps one helpful way of thinking about volition is as the executory mental state that produces an action in response to an intention.[78] But viewed in this way, virtually no one with a mental disorder has a volitional problem because people with disorders are fully able to execute the intentions that their disordered thoughts,

---

Mossman, but could make the interpretation narrower or even broader. Again, no term for the cognitive test is self-defining or talismanic.

[77] See, Stephen J. Morse, *Uncontrollable Urges and Irrational People,* 82 VA. L. REV. 1025 ,___ (2002).

[78] Among legal philosophers, this conception has been advanced best by Michael S. Moore, ACT AND CRIME 113-165 (1993).

perceptions and desires may motivate.[79] There may be some unusual cases when a symptom of mental disorder, such as anxiety, "paralyzes" a person who has a legal duty to act. In such instances, there is indeed a volitional problem, but such cases are rare to the vanishing point in criminal law. There is a common sense meaning of lack of self-control, of course, but terming it a volitional problem perpetuates a conceptual and pseudo-scientific confusion.

Despite the definitional and empirical difficulties, however, some criminologists and legal philosophers nevertheless believe control tests are necessary for crimes committed by defendants with impulse disorders, such as Intermittent Explosive Disorder,[80] or "disorders of desire," such as Pedophilia.[81] In the former case, however, the problem is that the person appears intermittently to be unable to bring rationality to bear, and thus such cases also appear best subsumed under a cognitive test. In the latter type of case, the defendant is usually entirely in touch with reality and probably knows and endorses the legal and moral code. Nonetheless, he claims that he cannot "help" or "control" his conduct.

The problem in cases of alleged lack of self-control is distinguishing the disordered person from anyone who also wants to do something very badly that they should not do, such as the very greedy person tempted terribly to steal. Simply characterizing the desire as disordered in one case but not the other cannot resolve the

---

[79] Herbert Fingarette & Anne Fingarette Haase, MENTAL DISABILITIES AND CRIMINAL RESPONSIBILITY 44-65 (1979).

[80] E.g. DSM-IV-TR, supra note XXX at 663-667. Recent data indicate that the 12 month prevalence of this disorder is 3.9% of the population and the lifetime prevalence is 7.3%. Ronald C. Kessler et al, *The Prevalence and Correlates of* DSM-IV *Intermittent Explosive Disorder in the National Comorbidity Survey Replication*, 63 ARCH. GEN. PSYCH. 669 (2006).

[81] Id. at 571-72.

question of control; it simply begs that question. Why should we ever excuse someone who acts wrongly in response to a very strong desire, whether that desire is normal or abnormal? Moreover, what theory or account allows us to characterize desires, as opposed to perceptions and beliefs, as irrational?[82] And, how do we distinguish between an irresistible desire and a desire simply not resisted?

Concerns like these led both the American Bar Association and the American Psychiatric Association to recommend abolition of an independent control test during the insanity reform movement that occurred in response to the Hinckley verdict.[83] One of us has argued that control cases can always be assimilated to cognitive cases in any situation in which an excuse is plausible. [84] It is sufficient here, however, to recognize that control tests are problematic and do not avoid the interpretive problems that beset more familiar cognitive tests.

The relevance of mental disorder to legal insanity tests is conceptually straightforward, although evidentiary problems can of course arise. Insanity tests address a broader issue than either claims concerning action or mens rea. Moreover, they are not tests about mechanisms in the literal sense. They do not address brain anatomy or physiology, for example. They do address acting human beings who may have difficulties with cognition or control. The finder of fact needs the thickest possible description of what the defendant was perceiving, thinking and feeling at the time of the

---

[82] Robert Nozick, THE NATURE OF RATIONALITY 139-40 (1993) (denying the existence of a substantive theory of the rationality of desires).

[83] AM. BAR ASS'N CRIMINAL JUSTICE STANDARDS COMM., ABA CRIMINAL JUSTICE MENTAL HEALTH STANDARDS 330, 339-42 (1989); AM. PSYCHIATRIC ASS'N, INSANITY DEFENSE POSITION STATEMENT (1982).

[84] Stephen J. Morse, note ___ supra.

crime in order to determine whether the legal standard for insanity is met. Mental disorder can affect perceptions, thoughts, beliefs, desires, and feelings, all of which need to be assessed to evaluate a claim for legal insanity under either a cognitive or control test. Anatomical, physiological and other kinds of non-behavioral evidence may help the finder of fact make inferences about the defendant's psychological phenomenology, but the ultimate test is behavioral, and mental disorder evidence is crucially relevant.

Before concluding the discussion of legal insanity, it is important to clarify a persistent confusion about the insanity defense: the claim that the criminal act was "caused" by mental disorder does not mean that the agent lacked free will or that the agent was determined. Moreover, free will or the absence of determinism is neither a necessary nor a sufficient condition of insanity under any definition. No insanity defense standard requires a finding that the defendant lacked "free will" or that his behavior was "determined." These locutions are often used to signal a conclusion that the defendant was or was not legally insane, but lack of free will or determinism is not a criterion for legal insanity and does no real explanatory work. People are found legally insane because they lack rational capacity or, more controversially, because they cannot conform their behavior to the requirements of law. Free will and determinism have nothing to do with this legal issue.[85] Failing to understand this is a ceaseless source of confusion about the need and criteria for an insanity defense.[86]

Further, the "causal role" that mental disorder may play in the criminal behavior must be properly understood. Causation in this instance means that mental disorder

[85] Stephen J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology,* LAW AND THE BEHAVIORAL SCIENCES (in press, 2007).
[86] See, e.g., Slobogin, note ** supra, at 1222, 1238 (confusing causes with reasons and implying that determinism plays an analytic role in insanity determinations).

produced distorted perceptions, thoughts or desires that influenced the defendant's reasons for action. It does not refer to "mechanical" causation. It has nothing to do with determinism or free will. The actions of people motivated in part by abnormal perceptions, beliefs and desires are actions, but they may be excused if the abnormality renders the agent sufficiently irrational. Causation of behavior by abnormal beliefs, for example, is no different from, no more "causal" than, causation of behavior by normal beliefs. The only difference is that in the former case, the agent may be irrational and should therefore be excused.

Lack of rational capacity—not determinism, lack of free will or abnormal causation--are what justify the insanity defense and explain its criteria.[87]

E. Mens Rea and Legal Insanity

As should be clear from the foregoing sections, 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 disorder are not automatons; rather they are agents who act for reasons. Their reasons may be motivated by distorted perceptions and beliefs, for example, 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, as Daniel M'Naghten and Andrea Yates presumably were.

In some rare cases, as we have already discussed, evidence of mental disorder might negate mens rea because the mental state it produces will be flatly inconsistent

[87] Determinism and causation are discussed further at text accompanying notes **--** infra.

with the mens rea required by the definition of the crime, or because it helps indirectly to explain why mens rea was not formed on that occasion.[88] In these cases, the same evidence that a defendant was delusional may both negate mens rea and support a finding of legal insanity, [89] but the questions being answered by the evidence are different.

IV. Clark v. Arizona[90]

A. Background

In the early morning of June 21, 2000, Eric Clark, a 17 year-old resident of Flagstaff, Arizona, was riding around in his pickup truck blaring loud music. Responding to complaints about the noise, Officer Jeffrey Moritz turned on the emergency lights and siren of his marked patrol car and pulled Clark over. Moritz left the patrol car and told Clark to remain where he was. Less than a minute later, Clark shot and killed Moritz. Clark was charged with intentionally killing a police officer knowing that the officer was acting in the line of duty.[91]

Clark did not contest the shooting and death. But he did claim he was suffering from paranoid schizophrenia at the time of the crime. He claimed both that his mental disorder negated the required mens rea for the crime charged (because he lacked the intent to kill a person and the knowledge that the victim was a police officer) and that in any event it rendered him legally insane.

---

[88] See text accompanying notes ** to ** supra.

[89] For example, if Clark really believed he was killing an alien impersonating a police officer, he both lacked knowledge that he was killing a police officer, thus negating the mens rea of knowledge, and for the same reason did not know what he was doing or that it was wrong.

[90] *Clark*, supra note ___. All future references to the opinion will simply cite the page number.

[91] All statements of the facts are taken from the Supreme Court opinion.

There was substantial evidence to suggest that Clark knew Moritz was a police officer and that he had planned just such a shooting, including Clark's statements to classmates a few weeks earlier that he wanted to shoot police officers. He had even arguably lured Officer Moritz by driving his truck with its radio blaring in a residential area.

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. These actions were plausibly a result of his paranoid delusions. Indeed, lay and expert testimony reported 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. The defense expert also testified that Clark had turned the radio up to drown out auditory hallucinations.

The operative Arizona legal insanity test under which Clark was tried was limited to the cognitive right/wrong test: once the burden of production was met, the defendant had the burden of proving by clear and convincing evidence that "at the time of the commission of the criminal act [he] was afflicted with a mental disease or defect of such severity that [he] did not know the criminal act was wrong."[92] The operative 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

---

[92] ARIZ. REV. STAT. ANN., Sec. 13-502(A) (West 2001).

was inadmissible to negate specific intent and that evidence of mental disorder short of legal insanity was not admissible to negate any *mens rea* element.[93]

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. He ultimately ruled, however, that *Mott* barred him from using it to consider mens rea and that he could consider this evidence only to decide the issue of legal insanity. He found Clark guilty of first degree murder, and, although specifically finding that Clark was suffering from paranoid schizophrenia and had distorted perceptions, nonetheless found him legally sane because he concluded that Clark knew his actions were wrong.

Clark appealed on the grounds that both the *Mott* rule and Arizona's narrow insanity defense (which, as mentioned above, included only one of the two disjunctive *M'Naghten* prongs[94]) violated due process. The Arizona court of appeals affirmed the conviction, the Arizona Supreme Court denied certiorari, and the United States Supreme Court granted certiorari to decide "whether due process prohibits Arizona from thus narrowing its insanity test or from excluding evidence of mental illness and incapacity due to mental illness to rebut evidence of the requisite criminal intent."[95]

B. The Insanity Issue

[93] 187 Ariz. 536,541, 931 P. 2d 1046, 1051 (en banc), cert. denied, 520 U.S. 1234 (1997). The Arizona Supreme Court mischaracterized the use of battered woman syndrome evidence offered to negate mens rea as a "diminished capacity" defense. It 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.

[94] See note ** supra and accompanying text.

[95] 546 U.S. ____ (2005).

Clark argued that the full *M'Naghten* rule was the minimum test necessary to satisfy due process and that Arizona's truncated test was therefore unconstitutional. Writing for the five-Justice majority,[96] Justice Souter rejected this claim, concluding that the full *M'Naghten* rule is not a fundamental principle of justice subsumed by the due process clause.

The Court correctly noted that the history of legal insanity defenses in this country demonstrates that there is substantial diversity of language and interpretation within the broad cognitive and control categories As Justice Souter pointed out, insanity definitions vary widely across the United States, and four states have abolished the defense entirely.[97] The Court also observed that the test for legal insanity is not a test for mental disorder. The tests for insanity and disorder have been devised for different purposes—assessing criminal responsibility and justifying mental health treatment—and there is controversy about both. This is inevitable because the test for legal insanity is a matter of policy. The Court concluded that because there is so much variation, "no particular formulation has evolved into a baseline for due process, and…the insanity rule, like the conceptualization of criminal offenses, is substantially open to state choice."[98]

Although reasonable people might believe that Arizona's truncated insanity test is not optimally just, the Court's holding on this issue—that due process does not require

---

[96] Chief Justice Roberts and Justices Scalia, Thomas and Alito joined Justice Souter's majority opinion. Justice Breyer concurred in the majority's analysis, but dissented from the result because he would have remanded the case to let the Arizona courts apply the new rule announced by the Court. Justice Kennedy wrote the dissent, in which Justices Stevens and Ginsburg joined.

[97] See note ** supra.

[98] P.___

"any single canonical formulation of legal insanity"[99]—seems plainly right, especially after *Mullaney* and *Egelhoff*.[100] But in the course of reaching this rather unremarkable conclusion, the Court may have confounded the two *M'Naghten* prongs.

In its original characterization of the disjunctive *M'Naghten* rule, as well as in its categorization of types of insanity rules, it referred to knowledge of the nature and quality of one's act as a question of "cognitive capacity" and knowledge of right and wrong as a question of "moral capacity."[101] But, in fact, both are cognitive questions, and indeed that is why both are usually alternative prongs of cognitive tests such as *M''Naghten*. They differ only in the object of the knowledge required. Moreover, as discussed above, mental disorder seldom disables a person's moral compass. The person may be making a "moral mistake" because his or her perceptions and beliefs are distorted by disorder, but the moral sense generally remains intact. Andrea Yates might genuinely have thought her behavior was morally justified, but her moral capacity was hardly disabled. Indeed, one could view her act as an indication that her moral sense was perfectly intact, albeit driven by a delusional belief.

The Court also consistently referred to control tests as "volitional" and characterized them as asking "whether a person was so lacking in volition due to a mental defect or illness that he could not have controlled his actions."[102] We recognize that it is common to refer to control tests as volitional, but, as we argue above, this is a confused

---

[99] P. ***.

[100] See text accompanying notes ** to ** supra.

[101] PP. *** and ***.

[102] P. ___

locution that should be abandoned.[103] Whether Clark thought he was killing an alien to save himself or thought he was killing an officer because he was angry at the police, his volition or will perfectly and competently executed the intention he formed. Yates perfectly executed her intention to kill, motivated by her desire to save the children from Satan's eternal torments. Both may have experienced grave difficulty conforming to the law because they suffered profound delusions about the nature of the world, but in neither case was a defect of the will the source of the problem.

The Court also decided that the narrow Arizona legal insanity rule was constitutionally acceptable because evidence of so-called cognitive incapacity—lack of knowledge of what one was doing—is relevant to the right/wrong test and has the same significance for both. The Court recognized that one might show lack of moral knowledge without showing lack of factual knowledge about what one was doing, but it correctly observed that lack of the latter was sufficient to show lack of moral knowledge.[104] After all, if the agent does not know what he is doing, he cannot know that it is right or wrong. Indeed, the Court interpreted Arizona's legislative narrowing of the rule as a "streamlining" change rather than a genuinely substantive alteration.[105] Even

---

[103] Infra Part _____.

[104] Contrary to the Court's assertion, it is quite possible that a defendant might not know what he or she was doing but would know that it was wrong. For example, suppose a defendant violently attacked a person with the belief that the person was a dog. See Joseph Livermore & Paul E. Meehl, *The Virtues of M'Naghten*, 51 MINN. L. REV. 769, 809 (1967). In such a case, the defendant would not know what he or she was doing, but would know that cruelty to animals was a moral and legal wrong. It is not clear from *Clark's* reasoning if a rule like Arizona's that produced such an outcome would be constitutional because the Court did not envisage this possibility. There is a serious question about whether such a rule would permit unjust blame and punishment. Nevertheless, although such cases are a theoretical possibility, they will be so rare that the Court's analysis is reasonable.

[105] P. ___, n. 24.

under the narrow rule, then, all evidence of lack of factual knowledge would apparently be relevant and admissible, as it was at trial.

C. The Mens Rea Issue

Clark's second due process claim was a challenge to the constitutionality of the Arizona rule, announced in *State v. Mott*,[106] that excluded expert mental health testimony on the issue of mens rea. In addressing this issue, the Court's majority goes quite wrong, not only confounding mens rea and insanity, but forcing psychiatric evidence into an arbitrary system of classification that is not only unworkable but that had never been considered or argued by Arizona.

Although the extent of the *Mott* rule's evidentiary exclusion is not entirely clear, it is clear that it prevented defendants from introducing substantial, relevant and reliable expert testimony concerning whether or not the requisite mens rea was formed in fact. Recall that Clark's factual claim was that, as a result of mental disorder, he believed that Officer Moritz was a dangerous "alien" impersonating a police officer. If this were true, then Clark was simply not guilty of intentionally killing a person—unless the definition of "person" were expanded to include extraterrestrials—let alone guilty of knowingly killing a police officer acting in the line of duty—unless the definition of "police officer in the line of duty" were expanded to include an extraterrestrial whose alien line of duty was to impersonate a police officer.

---

[106] Supra note ___.

The trial judge permitted Clark to introduce expert testimony that addressed the mens rea question,[107] but ultimately decided that *Mott* barred him from considering this testimony in reaching a verdict on the homicide charge. Thus, Clark was functionally prevented 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 majority mischaracterized Arizona's rule. It held that *Mott* did not really say what it said, and ruled instead that *Mott* actually applied only to two of the three types of mens rea evidence that the Court distinguished: observation evidence, mental disease evidence and capacity evidence. 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.

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."[108] 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.[109]

---

[107] A psychiatrist called by the defense not only testified that Clark was suffering from paranoid schizophrenia with delusions at the time of the killing, and was therefore insane, but also that because of those delusions Clark could not have formed the intent to kill Officer Moritz with knowledge that Moritz was a policeman. P.***

[108] Id. at **.

[109] Id. 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

The second category, "mental-disease evidence," the majority defined as, "opinion testimony that [a defendant] suffered from a mental disease with features described by the witness." That is, such testimony provides general information about the mental disorder from which the specific defendant allegedly suffers. The Court pronounced that this kind of evidence, which is almost always provided by mental health experts, inadmissible under *Mott.*

The majority defined the third category, "capacity evidence," as evidence "about a defendant's "capacity for cognition and moral judgment (and ultimately also his capacity to form *mens rea*)."[110] This, too, is opinion evidence provided by experts, and, as with category two, the Court upheld *Mott*'s exclusion of this type of mens rea evidence.

Rather astonishingly, 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."[111] Is this a typographical error, or a deeper reflection of the majority's own confusion between these two very different principles?[112]

---

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.

[110] Id. at ___(parenthetical in original).

[111] Id. at ***. [emphasis supplied].

[112] This is not an isolated example of the majority conflating insanity and mens rea. As discussed in the text accompanying notes ** to ** below, they do it again later in the opinion.

The majority's tripartite construction not only fails to do the theoretical work necessary to draw a sensible line between which types of expert mental disorder evidence the states may and may not exclude, it is wholly unprecedented. This classification was not part of Arizona law (or any state law we 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.

Moreover, this construction not only allowed the Court's majority to find the artificially restricted rule constitutional, it also allowed it, alternatively, to conclude that Clark failed to preserve his constitutional argument on this point because he never asked the trial court to admit mental state evidence in category one despite the exclusion of categories two and tree.[113] As Justice Kennedy aptly put it in his dissent, "Seizing upon a theory invented here by the Court itself, the Court narrows Clark's claim so he cannot raise the point everyone else thought was involved in the case."[114]

Justice Kennedy, and to some extent Justice Breyer in his 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. 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

[113] Id. at *** .

[114] Id. at *** (Kennedy, J., dissenting).

observations of others. That is to say, much of the inferential expert evidence Justice Souter believes *Mott* may constitutionally exclude is itself, at bottom, 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 a recognized diagnostic category that unifies those observations?

All clinical judgments are, by their very nature, informed not just by the clinician's observation of the particular patient being seen, but by the accumulated wisdom of observations of other patients by themselves and other clinicians and by the findings from empirical studies. This is what medicine, including psychiatry, is all about. The Court's fictitious categories of evidence not only bleed into one another, they seem peculiarly unable to do the important constitutional work the Court asks of them.

The Court also reasoned that Clark's due process claim depended on the application of the presumption of innocence, the presumption of sanity and "the principle that a criminal defendant is entitled to present relevant and favorable evidence on an element of the offense charged."[115] The Court noted, as it was bound to under *Winship*, that the presumption of innocence could only be overcome by proof beyond a reasonable doubt of all elements, including mens rea. This is boilerplate.

The Court then observed that the presumption of sanity is universal in some form, but then proceeded confusingly to mischaracterize 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."[116] Once again, this time in the heart of the opinion explaining

[115] Id. at___.

why Arizona may constitutionally prevent a defendant from introducing category two and three evidence to negate mens rea, the Court confuses mens rea with insanity. Indeed, the whole presumption-of-sanity discussion in this section of the opinion is not in the least relevant to Clark's mens rea claim.

The Court not only blurred the distinction between mens rea and insanity, it made wholly inconsistent observations about the relationship between the two. It rejected the argument that mens rea and insanity are "entirely distinguishable," yet then obscurely noted that insanity "trumps" mens rea, suggesting that they are not functionally distinguishable after all, because the former subsumes the latter when both are claimed.[117] The majority did correctly acknowledge 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.[118] Given that acknowledgment, and its obeisance to precedential platitudes about a defendant's constitutional right to present evidence that negates an element, how did the majority conclude that Arizona could constitutionally deprive Clark of that right? The majority reached this conclusion by claiming that Arizona could permissibly "channel" mental disorder evidence solely into the insanity issue and out of the mens rea issue because Arizona had legitimate state interests in such channeling.

2. The Channeling Argument

The Court accepted Clark's characterization of the *Mott* rule as a rule of evidence rather than as a reworking of the elements of homicide, and conceded again that

---

[116] Id. at___

[117] Id. at ***

the evidence was relevant. But, of course, as we have already seen in *Montana v. Egelhoff*,[119] states may preclude relevant defense evidence—even evidence that rebuts an element of the offense—if the state has a legitimate purpose in excluding the evidence. What were Arizona's "legitimate purposes" in channeling some forms of expert mental state evidence into insanity and away from mens rea? Here is where the Court's reasoning goes radically wrong.

The Court's first identified "legitimate reason" is yet another example of its conflation of mens rea and insanity. In the first part of the opinion, the Court had already reaffirmed Arizona's authority to define legal insanity as it wished and to place the burden of persuasion for insanity on the defendant. Consequently, the Court reasoned that if Arizona is to have this authority in practice as well as theory, it "must be able to deny a defendant the opportunity to displace the presumption of sanity more easily when addressing a different issue…."[120] This is a blinding non-sequitur, based on confusions about the presumption of sanity and "capacity" that we addressed above.[121] There is no presumption of sanity applicable to the mens rea elements.

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

---

[118] Id. at ***.

[119] See text accompanying notes ** to ** supra.

[120] P. ___

[121] See text accompanying notes ** to ** supra.

logical consequence of the interaction between the structure of criminal culpability and *Winship*'s constitutional requirements. Criminal responsibility can be avoided either by negating an element of the crime charged or by establishing an affirmative defense. Permitting an affirmative defense, as the Court recognizes, 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 articulated as a second "legitimate purpose" Arizona's desire "to avoid a second avenue for exploring capacity, less stringent for a defendant."[122] But mens rea is not about "capacity" or about "criminal responsibility" broadly; it is about whether a mental state required by the definition of the crime charged was in fact formed. 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 stand-alone defense."[123] Negation of mens rea is not an independent "defense" of "diminished capacity." Mens rea negation is is simply a straightforward denial of the prima facie case that needs no special name. The question of "diminished capacity" is the province of the insanity defense or mitigating evidence at sentencing, not mens rea negation. The state undeniably has the authority to reject a partial responsibility mitigating doctrine, but this was not what Clark was requesting. The majority's constant blurring of the mens rea and legal insanity issues simply perpetuates this confusion.

---

[122] Id. at ***.

[123] Id. at ___.

The Court then articulated a third legitimate purpose: Arizona has made a determination that mental state evidence is just too unreliable and risky to be the basis of a complete defense under which a defendant is entirely acquitted and unconditionally freed. Now, at last, we are at the transparent heart of Arizona's real concerns. In fact, the Court identified three separate such risks: 1) the controversial character of some categories of mental disease; 2) the potential of disease evidence to mislead; and 3) the danger of according greater certainty to capacity evidence than experts claim for it.[124] 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.

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.[125] As a result, caution is warranted "in treating psychological classifications as predicates for *excusing* otherwise criminal conduct."[126] 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.

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. But it is not a justification for channeling in the *Mott* direction, which *allows* excusing evidence but *disallows* evidence that would

---

[124] Id. at. ____

[125] See text accompanying notes ** to ** supra.

[126] P. ____ [emphasis supplied].

exonerate because it defeats the prima facie case. If Eric Clark genuinely thought that he was killing an alien impersonating an officer, he is simply not guilty, full stop, of homicide. He killed a person, but he did not commit the crime of homicide of a police officer, which requires that he intentionally kill a person with knowledge that the victim was a police officer.

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 heterogeneity in psychiatric diagnostic categories, and imperfect fit, as the American Psychiatric Association recognizes, between those categories and legal questions.[127] 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. Given the dangers of mental-disease evidence to mislead, the Court concluded, and we 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.

---

[127] *DSM-IV-TR*, supra note ***, at xxxii-xxxiii, xxxvii.

But this justification for complete channeling of expert evidence to the insanity issue proves too much. 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. As we have already discussed, states use diagnostic information in a host of other legal contexts.[128] 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, as we noted above,[129] 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 his cognitive and moral capacities, that is, about whether he was legally insane. But legal insanity is a legal question to be resolved by a lay jury or judge. Why should 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.[130] 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 and reliable evidence.

---

[128] See text accompanying notes ** to ** supra.
[129] See text accompany note ** supra.

[130] FED. R. EVID. 704(B).

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. 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. 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. 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. As we have discussed previously,[131] evidence about a defendant's capacity to form mens rea is extremely problematic and often lacks a solid clinical or scientific foundation. The criminal law would be better off if capacity evidence capacity were strictly limited or even prohibited. Again, however, the Court's argument proves too much, since capacity evidence, despite its limitations, is permitted to address every other question in criminal and civil law to which it may be relevant. Why permit its limitation here, when so much is at stake for the criminal defendant?

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. But the

---

[131] Part ____supra

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,[132] 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."[133] 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*,[134] which involved a due process challenge to the

---

[132] 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.

[133] P. ___

[134] 463 U.S. 880 (1983).

admission of a prosecution expert's psychiatric opinions about a defendant's future dangerousness. These opinions were elicited in the death phase of a capital trial, and were based entirely on the psychiatrist's response to the prosecutor's hypothetical questions. 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. 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.

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. Clark may or may not have believed that Officer Moritz was an alien—indeed, the trial judge concluded that he did not. But if the trial judge really did not consider, was able to erase from his judgment process, the relevant and reliable expert evidence that *Mott* excluded, Clark did not have a fair chance to cast reasonable doubt on the mens rea necessary to convict him of first degree murder.

V. The Constitutionality and Wisdom of Abolishing Mens Rea and Legal Insanity

We consider in this Part the question of whether due process mandates the requirement of mens rea for prima facie guilt and an insanity defense, and, even if the answer is "no," whether states should nevertheless retain these docrines. After all, if states are or should be free to abolish these doctrines entirely, defendants like Clark will be hard pressed to argue that states may not impose reasonable limitations on them, including truncating the *M'Naghten* test or channeling most expert mental abnormality evidence into insanity.

A. Abolishing Mens Rea

We will not linger over whether states may constitutionally abolish mens rea. This 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.[135] Nonetheless, it has been clear at least since *Morrisette v. United States*[136] that mens rea is "essential" to criminal law, and that states and Congress bear a heavy burden of justifying departures from it. But the Supreme Court has never crossed the line from "essential" to "constitutionally required," even for common law crimes at the core of criminal law whose mens rea elements predated the Constitution.

Part of the reason is that no legislature, to our knowledge, has ever been so bold as to purport to convert a serious common law crime into a strict liability crime.

[135] E.g., 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); United States v. Dotterweich, 320 U.S. 277 (upholding strict liability and punishment for up to a year for shipping misbranded pharmaceuticals).

Consequently, this difficult constitutional question has lingered at the edges of justiciability, drawing the considerable attention of academics but not courts.[137]

Clark is especially disappointing because, unlike any of the other handful of cases decided by the Court in this context, it was tailor-made for exposition about the constitutional pedigree of mens rea. The Court could have made short shrift of Clark's attack on the channeling rule by announcing that Arizona was constitutionally free to abolish mens rea entirely, and thus could constitutionally limit evidence that would negate mens rea. It could even have upheld the rule while acknowledging that Arizona was not free to abolish mens rea. Instead, it avoided the issue entirely, at considerable intellectual cost.[138]

We see no reason to think the Court will depart from the deference-to-states course charted in *Egelhoff* and *Clark.*[139] But to the extent the constitutional question requires a defense of mens rea as a core principle of blameworthiness, we gladly take up

---

[136] 342 U.S. 246 (1952).

[137] Compare Herbert L. Packer, Mens Rea and the Supreme Court, 1962 SUP. CT. REV. 107 (suggesting constitutional limits to the legislative abolition of mens rea) with Louis Bilionis, *Process, The Constitution, and Substantive Criminal Law,* 96 MICH. L. REV. 1269, 1278-79 (1998) (calling the notion that individual blameworthiness is central to the criminal law a "myth"). Complicating the debate is the impact of punishment theory. Some commentators have argued that a retributionist might, in some circumstances, care less about mens rea than a rehabilitationist. See, e.g., See Kenneth W. Simons, *When Is Strict Criminal Liability Just?* 87 J. Crim. L. & Criminology 1075 (1997). Though this detour is beyond the scope of this article, we must mention here that neither of us, who count ourselves as essentially in the retributionist camp, shares the idea that mens rea is separable from blameworthiness. On the contrary, as we discuss in the text accompanying notes ** to ** infra, the very idea of blameworthiness likely has its roots in our evolutionary ability to recognize intentionality.

[138] See text accompanying note ** to ** supra.

[139] See text accompanying notes ** to ** supra. Even after *Egelhoff*, some commentators remained hopeful that the Court might breathe constitutional life into mens rea. See, e.g., Richard Singer & Douglas Husak, Of Innocence and Innocents: The Supreme Court Since Herbert Packer, 2 BUFF. CRIM L. REV. 859 (1999), Alas, after *Clark*, we suggest those hopes are dashed.

the challenge. We do so despite enduring skepticism about substantive due process,[140] and about the constitutional authority of federal courts to interfere in such a fundamental and traditional state task as defining crimes. But if there is any such thing as a fundamental principle rooted in the common law, surely it must be the notion that common law crimes and other crimes implying moral blameworthiness require mens rea, and that the state may not exercise its awesome punitive power to punish mere accidents.

Indeed, robust doctrines of mens rea antedated the common law, and although there are of course cross-cultural and temporal variations, the notion that intentional acts deserve a different kind of response than accidental ones seems to be a human universal.[141] Part of the confusion about the historical-legal pedigree of mens rea is that the criminal law itself is a relatively recent invention. For most of civilization's history, the state simply did not concern itself with what we call "crime." With a few notable exceptions,[142] the ancient remedy for wrongs done by one person against another was

---

[140] See note ** supra.

[141] See Paul H. Robinson, A Brief History of Distinctions in Criminal Culpability, 31 HAST. L.J. 815, 996 (1980) (noting that "[w]hile Christian thought on mens rea had a dominant influence over its development in English law, similar concepts are found in nearly all criminal laws, often without a history of Christian influence").

[142] The Code of Hammurabi, for example, not only contained a comprehensive set of rules governing personal conduct, but made violators of some of those rules answerable to the state. See STANLEY A. COOK, THE LAWS OF MOSES AND THE CODE OF HAMMURABI 1-19 (1903). Its place in history was earned not only because of its scope, and its influence on western civilization through the Jews, but also because it was the first set of laws, and one of only a handful of known ancient laws, to make this remedial transition from private revenge to state-imposed punishment. Id. The Laws of Moses and the Justinian Code are two other famous examples of law codes that continued civilization's evolution from private revenge to state punishment. See COOK, supra; O.F. Robinson, Criminal Trials, in A COMPANION TO JUSTINIAN'S INSTITUTES (Ernest Metzger ed., 1998). Another example were Draco's laws from fifth century Athens, which, in addition to their famous severity, were the first known set of Greek laws to make homicide an offense punishable by the city-state, apparently in an effort to curb a rise in Athenian revenge killings. See DOUGLAS M. MACDOWELL, THE LAW IN CLASSICAL ATHENS42-43(1978).

almost exclusively a matter of private revenge.[143] Only a small number of violations--such as treason or other wrongs directed against the state itself—justified the attention of government.[144] Therefore, it is quite misleading for commentators trying to deny the pedigree of mens rea to argue that it is an invention of common law England,[145] when in fact that is when the English criminal law itself—not mens rea—began emerging.[146]

But our plea 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 simply immoral for the state to punish accidents and intentional harms the same way. The mens rea an offender possesses expresses his or her attitudes towards the rights and interests of fellow citizens.[147] Different culpable mental states justify different levels of punishment consistent with different degrees of desert. And without mens rea, no blame and punishment is justified at all, at least for core crimes. The causal roots of that immorality may be in part biological,[148] but however our concern with mens rea began, it is now a fundamental feature of our moral practices. One need not be a neo-Darwinist to

---

[143] Or to be more precise, our modern distinctions between public and private , between crime and tort, are just that--modern distinctions. See, e.g., James Lindgren, *Why the Ancients May Not Have Needed a System of Criminal Law,* 76 B.U. L. REV. 29 (1996*).* It is rather remarkable to contemplate that trial by battle, which was a formalized kind of private revenge, was not officially abolished in England until 1819, though it had, admittedly, all but disappeared by the end of the reign of Edward III. Edward J. White, LEGAL ANTIQUITIES 118 (1913).

[144] Lindgren, supra note 126, at 39. Justinian's Code, for example, recognized only a few crimes against the Roman public, including treason, adultery, assassination and parricide. Id.

[145] See, e.g., Deborah W. Denno, Criminal Law in a Post-Freudian World, 2005 U. ILL. L. REV. 601, 610 at n. 58, citing 2 Frederick Pollock & Frederic William Matiland, The History of English Law: Before The Time of Edward I, at 470-73 (2d ed. 1968), for the proposition that crimes were strict liability crimes in England until the fifth century.

[146] For the evolutionary roots of mens rea we need to look much further back than the emergence of punishing states, or even the existence of states at all. We look at those roots in Part ***.

[147] See Part XXX infra.

[148] We address this speculation in Part XXX infra.

understand that if "fundamental rights" or "self-evident principles" mean anything, they must include guarantees that the state must satisfy its duties under the consent-to-be-governed bargain. Neither we nor our founding or framing ancestors reasonably expected the state to be able to punish us for accidents, any more than Holmes' dog mistook kicks for stumbles.

Mens rea is so crucial to fair ascriptions of blame and imposition of punishment that we are wary of attempts to dilute it by redefinitions of elements and affirmative offenses or by evidentiary rules. Although the criminal law, as a human institution, 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, thus risking wrongful conviction or wrongful conviction of a more serious crime, unless there are supremely good reasons for doing so. The presumption should be against redefinitions and evidentiary rules that undermine the values *Winship* protects.

There are occasions when making a claim part of affirmative defenses rather than an element may be warranted because the claim really is an excuse, or because there are good evidentiary reasons for shifting the burden of persuasion to the defendant. For example, New York's "extreme emotional disturbance" doctrine addressed in *Patterson* is clearly a generic partial excuse and has no necessary logical or empirical connection with homicide. Likewise, one could argue that dissociated, act-like behaviors, such as sleepwalking, should be treated as affirmative defenses rather than as part of the definition of the act element because establishing whether such a state existed at the time of the crime can be extremely problematic. Treating such cases of automatism as an affirmative defense and shifting the burden of persuasion to the defendant, who is in the

best position to provide the evidence, might render verdicts more accurate in such cases. Again, however, we should be wary about such rules.

B. The Constitutionality and Wisdom of Abolishing the Insanity Defense

The authors of this article disagree about the strength of the justification for the defense of legal insanity, but we agree about the wisdom of retaining some form of the defense.[149] But first things first. May a state constitutionally abolish all forms of the insanity defense? Probably.

The Supreme Court noted in *Clark* that it has never ruled whether an affirmative defense of legal insanity is required by due process. A few older state court opinions held abolition unconstitutional as a violation of fundamental fairness and the right to a jury trial.[150] But more recently, five states have abolished the insanity defense, and in four the abolition was upheld by the state supreme court.[151] In the fifth, Nevada, the state supreme court held that abolition violated due process under both the state and federal constitution.[152] In short, the constitutionality of abolishing the insanity defense is largely an open question. How should it be resolved?

---

[149] One of us (Morse) is a strong supporter of retaining an insanity defense in some form; one (Hoffman) is less enthusiastic. We agree, however, that the defense should be retained, even if the constitution does not require it, and that it should be defined relatively narrowly. See text accompanying notes ** to ** infra.

[150] State v. Strasburg, 110 P. 1020 (Wash. 1910); Sinclair v. State, 132 So. 581 (Miss. 1931)

[151] State v. Korell, 690 P.2d 992 (Mont. 1984); State v. Searcy, 798 P. 2d 914 (Idaho 1990); State v. Herrera, 895 P.2d 359 (Utah 1995); State v. Bethel, 66 P.3d 840 (Kan. 2003).

[152] Finger v. State, 27 P.3d 66 (Nev. 2001).

Insanity as an excusing condition has as formidable a history as mens rea.[153] Forms of a recognizable insanity defense were established in antiquity, Hale recognized the importance of legal insanity in the 17th C., and when the House of Lords adopted the M'Naghten rules they were simply codifying what was already understood to be the common law.[154] Some form of the insanity defense was almost certainly part of the common law at the time of the founding and was in effect in all states at the time of the adoption of the 14th Amendment.. After M'Naghten, some type of insanity defense was universal until the 1980s. This suggests at least a substantially deep tradition in Anglo-American common law and in American statutes that followed the common law. Perhaps the insanity defense is so deeply rooted in our tradition that the constitution requires its retention in some form.

On the other hand, despite the similarities among the states' different approaches, there has often been profound disagreement, and even perplexity, about the justification for the excuse, and as a result long-standing inquiries into cognitive and control functions have generated varied approaches across time and jurisdictions. From time to time, as in the wake of the unpopular Hinckley verdict, there has been intense criticism of the defense.[155] Even if the brief post-Hinckley wave of abolition does not represent a discernable trend, perhaps there is no better evidence that the insanity defense is not

---

[153] A. Platt & B. Diamond, *The Origins of the "Right and Wrong" Test of Criminal Responsibility and Its Subsequent Development in the United States: An Historical Survey,* 54 CAL. L. REV. 1227 (1966); Daniel N. Robinson, WILD BEASTS & IDLE HUMOURS: THE INSANITY DEFENSE FROM ANTIQUITY TO THE PRESENT (1996).

[154] Platt & Diamond, note *** supra. Moreover, a famous earlier case had adopted the "product" test. Hadfield's Case, 27 Howell 1281 (1800). Isaac Ray's influential 1838 treatise is also indicative of the common law lineage of legal insanity. Isaac Ray, A TREATISE ON INSANITY (1838).

[155] The alleged empirical, historical and consequential reasons for dissatisfaction with insanity defense at the time were either empirically incorrect or analytically unsound. None was persuasive. .. Stephen J. Morse, *Excusing the Crazy: The Insanity Defense Reconsidered*, 58 S. CAL. L. REV. 777, 788-801 (1985). .

deeply rooted than the fact that five state legislatures have in fact abolished it.[156] Moreover, the language in *Powell* about the constantly shifting tension in the criminal law's treatment of responsibility,[157]suggests that the Court does not believe our core traditions require the insanity defense, although that language could also be interpreted to mean that no specific form of the insanity defense is required, as *Clark* held.

In any event, the constitutional pedigree of the insanity defense is an open question, and there is no indication that the current Court will reject the plausible argument that it is not constitutionally required. In the balance of this section, we argue that the insanity defense should continue to be an essential feature of any criminal system that claims to provide justice based on desert and blame. Indeed, we believe a robust but narrowly-defined insanity defense, and other appropriate excusing conditions, will paradoxically help enhance the importance of mens rea, and the understanding that almost all of us are sufficiently responsible for our actions to deserve blame and punishment if we offend.

The crucial normative question is whether it is unfair to blame and punish a defendant who meets the definitional elements of the crime, but whose rationality at the time of the offense was severely impaired.[158] Rationality is the touchstone of responsibility, as the structure of criminal law itself indicates. All laws, criminal and

---

[156] The weakness of the abolitionist arguments had been fully exposed by the time the "wavelet" of abolition occurred. The point, however, is that the very fact that five states have abolished the defense, and four states' highest courts have upheld abolition, might be a clue about whether the insanity defense is constitutionally imbedded in our deepest due process traditions. We recognize the tautological difficulties with this argument. This difficulty is evidence of the problematic nature of substantive due process.

[157] Supra note ____.

[158] We refer to impairment of rational capacity generally to avoid begging the question in favor of any particular legal insanity rule and because the dominant insanity rules involve rationality defects.

civil, make sense and are functional precisely because they are action-guiding reasons addressed to potentially rational creatures. We do not expect other animals to understand the reason for a rule or the deterrent value of punishment because other animals are not capable of the same degree of rationality as homo sapiens. There are no chimp legislatures or avian police. It is simply unfair to hold responsible, to blame and punish, mentally disordered wrongdoers who are not morally responsible because they were not capable of being rational at the time of the crime.

Consider, as an analogy, the criminal responsibility of juveniles. The common law defense of infancy conclusively presumes that children younger than seven are not responsible and imposes a rebuttable presumption of lack of responsibility for those between the ages of seven and fourteen.[159] The reason for this rule, of course, is that the rational capacities of young children are not formed, even though they may be entirely capable of forming and acting on intentions, and even though they might know that what they are doing is wrong. A four-year old is perfectly capable of intending to pocket a stick of gum, and of doing so even after his or her parents have said that the child may not have it. We do not prosecute the four-year old because we understand that four-year olds have just begun to internalize social norms and rules, and do not fully appreciate what it is about the world that requires items in stores to be paid for, or indeed, even what it means to "pay" for something.[160]

---

[159] LaFave, note *** supra, at 485-86.

[160] That we do not allow the state to punish them does not of course mean they should not be punished as a method of instruction unrelated to genuine moral desert. The whole process of internalizing the social contract is a learning process that requires instruction and example from the moment the infant begins to realize he or she is not the center of the universe. But to subject individuals to *state* sanction, the intentional infliction of stigma and pain, is, at least in enlightened societies, quite a different matter that requires desert.

More controversially, but driven by the same principles, the Supreme Court recently held that states may not apply their death penalties to criminally responsible capital murderers who suffered from retardation or were under the age of eighteen when they killed. [161] In both cases, the Court pointed to factors related to the general capacity for rationality, such as the retarded person's "diminished capacities to understand and to process information" and "to engage in logical reasoning"[162] and to the adolescent's "lack of maturity and an underdeveloped sense of responsibility."[163] The Court's reasoning confirms that lesser rational capacity, whether arising from normal developmental processes or abnormalities, is the reason adolescents and those with retardation are less responsible.

Of course, the capacity for rationality is clearly a continuum, ranging from almost none for two year olds or people with profound developmental disability or severe dementia, to fully intact for most adults with normal cognitive development. Despite conceptual and empirical disputes about mental disorder, it is clear that a handful of severe mental disorders, at the extreme of the continuum of disorders, can deprive people of the capacity to act rationally. For example, a person who delusionally believes that he is being threatened with imminent death is incapable of rationally considering whether he should act in self-defense. He is "rational" only in a narrow, instrumental, sense. If his psychotic premises were true, it would indeed be rational to act in self-defense, and in so

---

[161] Atkins v. Virginia, 536 U.S. 304, 318-319 (2002); Roper v. Simmons, 543 U.S. 551, 569 (2005). We may not necessarily agree with the Court's constitutional analyses in these cases, but the majority opinions in both relied heavily on behavioral science that indicates rationality differences between late adolescents and people with retardation compared to normal adults.

[162] *Atkins*, at 318;

[163] *Roper*, at 569.

acting he would no doubt be acting with the required mens rea. But such a narrow view of responsibility is conceptually and morally impoverished. Our hypothetical self-defender is not just making a mistake he could correct by being more careful. He is incapable of making a rational judgment about how to respond to the real world because he has grossly lost touch with that world.[164]

We share critics' concerns about the reliability of insanity evidence, and indeed have made the case here[165] and elsewhere[166] that the insanity defense's biggest problem is that it has been "over-scienced." In the end, unlike the mens rea inquiry, legal insanity is a legal and moral policy judgment, not a particular empirical fact. But our justified skepticism about the normative and even descriptive powers of psychiatry and psychology should not lead us wrongheadedly to throw the moral and legal baby out with the pseudo-scientific bathwater. Indeed, precisely because insanity is a moral construct and not a scientific one, we cannot abandon it to the scientists, especially scientists who, as we discuss below,[167] misguidedly reject the possibility of individual responsibility *tout court*. Science may, and should, inform our normative judgments about responsibility by

---

[164] Accord, Westen, note ** supra, at 154-55. Christopher Slogobin's "integrationist" alternative to the insanity defense, discussed further at text accompanying note ** supra, fails to make this distinction, in part because it misunderstands the relation between negligence and mental disorder. See text accompanying notes **-** supra and see Slobogin, *An End to Insanity*, note ** supra, at 1202,1240, and MINDING JUSTICE at 51-52, 56. Providing an excuse for any defendant who would be justified if the facts he believes were true would treat the careless and crazy self-defender alike. This is normatively undesirable because the crazy defender is an irrational and therefore non-responsible agent, whereas the careless defender has failed to exercise capacities he or she possesses and thus culpably violates the rights of his or her victim. The law convicts negligent self-defenders for murder or, in more forgiving states, of voluntary manslaughter based on "imperfect self defense." It properly does not acquit careless self-defenders.

[165] See text accompanying notes ** to ** supra.

[166] Morris B. Hoffman & Stephen J. Morse, "The Insanity Defense Goes Back on Trial," New York Times, July 31, 2006; Stephen J. Morse, *Crazy Behavior, Morals and Science: An Analysis of Mental Health Law*, 51 S. CAL. L. REV. 527 (1978); *Crazy Reasons*, 10 LAW & CONTEMP. PROBLEMS 189 (1999).

[167] See text accompany notes **-** infra.

providing the best possible understanding of human behavior. But it cannot and should not displace them.

How many people with mental disorders should be excused depends entirely on how broadly or narrowly the insanity defense is defined and construed. But no matter how limited an insanity defense may be, some defendants will have been so grossly out of touch with reality at the time of the crime that an excuse is warranted because their mental disorder, as in our example of the deluded self-defender, interferes with their rationality even if it does not interfere with their formation of mens rea. It is precisely because of this category of defendants that mens rea alone cannot do the work necessary for a morally coherent system of criminal responsibility.

Note that if Clark were believed, he would be the rarest of the rare, a defendant whose severe mental disorder produced a belief entirely inconsistent with mens rea on that occasion and undermined his general capacity rationality. But note also that without an insanity defense, Clark could still be convicted of "imperfect self-defense"/voluntary manslaughter or negligent homicide, since of course his belief that Officer Moritz was an alien was unreasonable. This is another important reason mens rea alone is not enough—because its broadest net (negligence) will catch people like Clark who are not at all being "negligent" in the ordinary sense. They are incapable of behaving more carefully because of their deluded premises, and are thus as arguably morally innocent as our four-year old. It is simply unjust to blame and punish them.

If successful insanity defenses, like those of Daniel M'Naghten and Andrea Yates, are so infrequent, why should states even bother with adjusting their responsibility

regimens to accommodate them?[168] The answer is that desert is a necessary justification of punishment, and the very nature of desert and responsibility is individual. No one should be blamed and punished unless they really deserve it, especially for core crimes.[169] With only a very few doctrinal exceptions, we hold a defendant responsible for what he and he alone did, and for what was in his mind, and his mind alone, at the time of the doing.[170] We do not generally attribute one individual's actions or intentions to others. Likewise, we will consider defenses of excuse only by applying the tests for a particular class of excuse to an individual case. Paranoid schizophrenics, as a group, are of course not necessarily excused from all criminal responsibility. But neither should they, as a group, necessarily be deemed always responsible.

As long as we agree that there might be some individuals who are perfectly capable of executing rational intentions based on a profoundly irrational, mentally disordered loss of reality, mens rea alone will never be enough. Our legal system has a moral duty to identify those people and to let jurors consider whether their crimes should be excused.[171]

---

[168] Morse, note *** supra, at

[169] C.S. Lewis made this point elegantly:

> What can be more immoral than to inflict suffering on me for the sake of deterring others if I do not deserve it? And what can be more outrageous than to catch me and submit me to a disagreeable process of moral improvement without my consent, unless (once more) I deserve it?

C.S. Lewis, THE PROBLEM OF PAIN (91-92) (1940).

[170] Genuine non-culpable ignorance of law, conspiracy, and felony-murder are counter-examples that stretch this principle of individual responsibility, but these doctrines are justified on purely consequential grounds and are considered exceptions to the principle in the text.

[171] We are not suggesting, for example, that all religious extremists who crash airplanes into buildings should necessarily be excused from their behavior because their extremism is a form of "loss of reality." See Stephen J. Morse, *The Jurisprudence of Craziness,* in F. Parisi & V. Smith, Eds, THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR 225, 256-258 (2005) (considering the distinction between irrational

D. Alternatives to Abolition

The preceding part of this article made clear why the first alternative, the mens rea approach that limits admission of mental abnormality evidence solely to the issue of mens rea formation, is unfair. In this section, we address three other alternatives to the insanity defense: using mental abnormality evidence at sentencing, the Guilty But Mentally Ill verdict, and a recent academic proposal for an "integrationist" approach to abolition that bears much similarity to the mens rea approach, We reject all three.

A second alternative to abolishing legal insanity as an excusing defense is to limit consideration of mental disorders to sentencing. This is a kind of sentencing version of the evidentiary channeling critiqued above.[172] But this alternative blurs the distinction between crime and punishment, and thus between blameworthiness and blame. Excusing conditions like insanity are excusing conditions precisely because they are deemed inconsistent with blameworthiness. They are defenses to the crime itself, not mitigating facts. Convicting morally innocent agents—such as the Model Penal Code's lemon squeezer[173]—cannot be acceptable in any system of criminal responsibility that claims to be based on individual blameworthiness.

Moreover, moving mental state issues from affirmative defenses to sentencing changes the decision-maker from jury to judge, at least in the vast majority of states and the federal system, where juries play no role in sentencing. Yet, as we have argued

---

and rational terrorists). Here again, the command of a just rule of law is to consider cases individually, to allow individual members of such groups the chance to prove they are suffering from a recognized mental disease so profound that they lost touch with reality, and then let reasonable jurors decide. There is much evil in the world whose name is not insanity. But there is some whose is.

[172] See text accompanying note ** to ** supra.

[173] See text accompanying notes ** to ** supra.

throughout this article, the insanity inquiry is not scientific, or even primarily factual; it is a normative question about an individual's rational capacity and his resultant responsibility for breach of the social contract. This is a moral question best answered by ordinary people, not judges or mental health experts. Finally, because sentencing is largely discretionary, there is no guarantee that the appropriate role of mental abnormality will be considered or that individual sentencing judges will calibrate punishment properly even if they do take mental abnormality into account. Juries may not always make wise decisions in legal insanity cases, but the existence of the insanity defense will at least guarantee that mental abnormality is considered in appropriate cases.

The third modern alternative to abolishing the insanity defense is the "guilty but mentally ill" verdict ("GBMI"), which has been adopted by a substantial minority of states.[174] Unlike the sentencing alternative, which considers potential diminished responsibility after the defendant has been authoritatively labeled blameworthy and responsible by a guilty verdict, GBMI has nothing to do with responsibility. In fact, the convicted GBMI defendant is found fully culpable and there is no reduction in sentence.[175] A GBMI capital murderer can be executed.[176] Moreover, GBMI is not a treatment tool, since mental health evaluations and treatment are routine parts of health care in every prison system. Besides, when a GBMI convict is hospitalized and successfully treated, he is then returned to prison to complete his sentence, just like any other convict.

---

[174] *Clark* at 2721.

[175] Ralph Reisner, Christopher Slobogin & Arti Rai, LAW AND THE MENTAL HEALTH SYSTEM: CIVIL AND CRIMINAL ASPECTS (4TH ED.) 610-11 (2004).

[176] People v. Crews, 522 N.E. 2d 1167 (1988).

In short, GBMI is a politically expedient "third-way" fraud. It has nothing to do with responsibility and nothing to do with treatment. It is the equivalent of a verdict of "guilty but herpes," with no guarantee that the herpes will be treated or even noticed once the defendant leaves the courtroom. Its only purpose is to give legislatures political cover, and to lull jurors into mistakenly believing their GBMI verdicts might mean that defendants with mental disorder who receive this verdict will be given special treatment and attention, or otherwise dealt with more mercifully at sentencing or in prison.

The fourth and final alternative we wish to consider is Professor Christopher Slobogin's proposed "integrationist" alternative to the insanity defense, which is the only reasoned contemporary scholarly proposal for abolition that we are aware of.[177]

Professor Slobogin provides a critique of the present insanity defense as one ground of justification for his abolitionist proposal. Before turning to the proposal itself, we first present only a limited discussion of that critique because we believe that our positive defense of the insanity defense *supra* and our criticisms of the integrationist proposal expose the flaws.

Much of the integrationist critique of the present insanity defense relies on setting up straw men, on confusing mechanical causation and practical reason, and on weak analogies. For example, Professor Slobogin criticizes the "volitional" test by characterizing it as if causation and lack of control were equivalent, but this is a caricature of an adequate control test, rather than an attempt to come to grips with its best

[177] See, Christopher Slobogin, *Ending Insanity*, note ** supra, and MINDING JUSTICE, *id.*[hereinafter, MJ]. Although Professor Slobogin's argument does not change, we will focus here on the most recent, somewhat expanded version in the latter.

form and justification.[178] He criticizes an appreciation test for insanity by confusing statistical predictability with a practical reasoning excuse.[179]

Professor Slobogin criticizes the argument for a rationality-based justification (including one based on "appreciation") by noting that a rationality test might excuse psychopaths or pedophiles.[180] Such people are of course not popular, and public opinion and current law do not excuse them. But, once again, he fails to confront a genuine practical reasoning approach to these conditions, and there are very good arguments for questioning the responsibility of some people with these conditions on grounds of impaired rationality.[181] Professor's Slobogin's criticism here borrows from his criticism of "volitional" standards, but the argument depends on a misunderstanding of volitional problems and the confusion of causes and reasons. Virtually all philosophers who have addressed the issue argue that psychopaths are not morally responsible. Legal and moral responsibility may be treated differently, but if one believes that the law should not blame and punish those who are not morally responsible, there is a strong case for excusing some psychopaths. Indeed, Professor Slobogin is internally inconsistent because he earlier suggests that people who cannot understand the concepts of good and evil [i.e., psychopaths] would be excused under his proposal.[182]

---

[178] MJ at 39-41

[179] Id at 44 (noting that people with psychoses are less dangerous than is usually believed).

[180] Id at 42, 51.

[181] See, Stephen J. Morse, *Addiction, Genetics and Criminal Responsibility*, 69 LAW & CONTEMP. PROB. 160, 190-95 (2006) (providing such an account for addicts, but ultimately rejecting an excuse in most cases because most addicts can be held diachronously responsible); Susan Wolf, FREEDOM WITHIN REASON 120-121 (1990) (arguing that those who lack the ability to be receptive to the Good—e.g., some psychopaths—are criminally insane). SJM believes that severe psychopathy should qualify for an excuse (and post-acquittal commitment); MBH is not sure.

[182] MJ at 27.

Professor Slobogin appears in some places to understand that determinism and "mechanical" causation by factors outside one's control are not the justifications for legal insanity,[183] ., but his discussion often conflates these alleged justifications with genuine justifications.[184] All such arguments necessarily must fail analytically, so we shall focus in the text only briefly on those that may be analytically and normatively sound.

Here is the integrationist proposal:

> A person shall be excused from an offense if at the time of the offense [by reason of mental disease or defect] he (a) lacked the subjective mental state of the conduct circumstances, or result element of the crime; (b) believed circumstances existed that, if true, would have justified the offense; (c) believed circumstances existed that, if true, would have amounted to duress; or (d) was unaware of the general prohibitions of the criminal law…[185]

Slobogin essentially proposes that the same defenses that apply to other criminal defendants would be applied to defendants with mental illness and there would be no independent insanity defense. He places mental disease and defect in brackets to indicate that the language need not be part of the test. Indeed, the logic of his proposal is that it should not.

Professor Slobogin quite rightly claims that many people with mental disorder are aware of society's normative judgments about right and wrong conduct and that those who would be excused would be doing the right thing if the circumstances they believed were true. In contrast, he accuses the dominant formulations of the insanity defense of being "overbroad"[186] and unable to identify accurately those defendants who are

---

[183] *Id.* at 40

[184] *Id.* at 39.

[185] *Id.* at 56. Professor Slobogin would deny the defendant the defense if the defendant was responsible for having the erroneous beliefs because the person purposely avoided treatment with the awareness that such beliefs would occur without treatment. *Id.* We largely agree that defendants who culpably create their own excusing condition should not be excused.

[186] Id. at 24.

sufficiently rational to be excused,[187] apparently because the tests can be interpreted to excuse more people than he believes normatively deserve an excuse, of being "whimsical" because they allegedly permit the defendant to define the scope of the excuse[188] and because the justification for convicting mentally disordered people with lesser rationality problems but acquitting those with severe rationality problems is "often weak."[189] He wonders why irrationality is an excusing condition at all.[190]

Slobogin's criticisms betray a misunderstanding of the justifications for this standard and for the insanity defense in general. His argument about overbreadth confuses causation with excuse. All behavior is caused by the necessary and sufficient variables that produce it. Other factors, such as poverty, may play a stronger causal role in crime than mental disorder, but causation per se is not an excusing condition. Lack of rational capacity, not causation by some factor, is why people are excused and thus there is no danger of overbreadth by excusing defendants whose criminal behavior was produced by more causally powerful variables than mental disorder. Failure to understand this point was the central flaw in prior proposals for abolition in the wake of the Hinckley acquittal.[191]

Moreover, how broad or narrow an excuse should be is a central normative question. It is not a valid criticism of any insanity formulation to note (correctly) that it

---

[187] E.g., *id.* at 50.

[188] *Id.*

[189] *Id.* at 38.

[190] *Id.* at 46.

[191] E.g., Norval Morris, MADNESS AND THE CRIMINAL LAW 61-63 (1982). Professor Morris' argument heavily influenced the widely noted by equally misguided abolitionist proposal by the American Medical Association. Board of Trustees, *Insanity Defense in Criminal Trials and Limitations of Psychiatric*

may be interpreted too broadly for a particular critic's taste. All tests can be misinterpreted, and one's position on the normative issue of criminal justification requires a normative construction, not a prescriptive broadside. .We have already explained why the nature of law itself supports the view that rationality must be the primary responsibility condition and irrationality the primary excusing condition in both criminal and civil law.[192] Slobogin does not embed his critique in a concept of the person or a jurisprudence of responsibility and gives us no moral reason to prefer his outcome. There are line drawing problems with any test, including his own.[193] Identifying the right people to excuse as legally insane is an admittedly difficult enterprise, but it is no more difficult than many other line-drawing decisions the criminal law must make.[194]

Excusing a defendant as legally insane does not permit the defendant to define the scope of the excuse. The finder of fact must decide if the defendant's rationality was sufficiently impaired to meet the standard for the insanity defense.

Analogies between those suffering from severe cognitive abnormalities and others who do not, such as gang members or those with dependent personalities,[195] are neither

---

Testimony, 251 J.A.M.A. 2967 (1984). Both were criticized on these grounds. Stephen J. Morse, *Excusing the Crazy: The Insanity Defense Reconsidered*, 58 S.CAL L. REV. 777, 788-97 (1985).

[192] See text accompany note ** supra.

[193] For example, suppose, in the absence of a threat for failure to comply, a defendant hallucinates God's command to do a criminal deed or delusionally believes that God wishes him or her to commit the crime. If these facts were true, what would be the legal outcome? Justification, excuse, or neither? Cf. MJ, 27. Or, suppose a defendant delusionally believes that his life is in deadly danger, but, unknown to the defendant, he is about to be the victim of his believed aggressor and would have had a right to defend himself if he had actually known the facts. Suppose such a defendant were not deluded. Should he be justified, excused or neither?

[194] Professor Slobogin's attempt to analogize the difficulties besetting rationality and volitional assessments is marred by a failure to understand the difference between the nature of the cognitive and volitional phenomena being assessed and by a confusion between mechanical causation and causation by practical reason. E.g., MJ, at 47, 50.

[195] Id. at 50.

apt nor useful. All people, including people with severe mental disorders, have reasons for doing what they do, Likewise, all people can have their reason compromised, by anything from fleeting emotions to profound and persistent delusions. At some point, sufficient irrationality deprives people of moral agency. Of course, as Slobogin notes, if irrationality does not sufficiently affect the agent's reason for committing the crime, no excuse is warranted,[196] but the irrationality standard best explains why this is so,

The virtue of the irrationality standard is that it provides an anchor for a meaningful normative debate, and avoids the confusion of thinking that causation or determinism is the basis of the excuse or that causation is the equivalent of compulsion. Slobogin's integrationist proposal, in practice, will not provide the proper terms for evaluating blameworthiness and will not produce justice.[197]

Professor Slobogin believes that the doctrines of justification and negligence have largely been subjectivized, and proposes that defendants whose wrong beliefs about reality arise from psychotic states should be treated the same as those whose errors result from culpable (unreasonable) carelessness or inattention. He is incorrect about the extent of subjectivization that the Model Penal Code accepts,[198] and in fact his proposal would lead to the loss of the objective standard for both justification generally and negligence in particular. Slobogin is perfectly entitled to propose and argue for the desirability of such a radical change in our law, but he does not provide an argument for thoroughgoing subjectivization of the criteria for justification and negligence.

---

[196] Id.

[197] We do agree for the reasons given in Part.** supra that defendants should be permitted to use evidence of mental abnormality to negate mens rea and that full negation will seldom occur. Our disagreement and discussion will be limited to the affirmative defense parts of the proposal.

To exonerate a culpably careless wrongdoer the same as a deluded wrongdoer makes a mockery of the moral differences between them. The culpably careless have the capacity to pay attention and to be as careful as we expect them to be, yet they have not exercised that capacity. They have failed to respect the rights and interests of their victims. Such failure is the essence of blameworthiness in our system of morality and law. Delusional defendants, on the other hand, have not simply made a mistake that could have been corrected by being more careful. They are incapable of getting the facts right because they are deluded. They are quintessentially blameless in our system of morality and law.[199]

The integrationist approach also does not have the resources adequately to address the familiar case of psychotic mental states with religious content, such as hallucinating that one is hearing God's voice or delusionally believing that God wants the person to behave in some way. Suppose that there is no apparent threat to the agent, who simply believes that he is doing God's will. Such a person would not be justified under traditional criminal law, even if the facts and circumstances were as the agent believed. The standard for the balance of evils portion of the residual "choice of evils" justification is purely objective,[200] and the agent may be convicted of a crime of negligence for negligently misappraising the situation. There is no subjectivization of the reasonable person standard.[201]

But some of these agents act for such irrational reasons that it would be unjust to blame and punish them. Of course it can be difficult to determine whether a person

---

[198] See note ** supra.

[199] Indeed, they would be excused under the "insane delusion" part of the M'Naghten rules, so the integrationist approach is not needed for these defendants.

motivated by religious beliefs is psychotic, but this is no reason not to try to achieve justice. Nitpicking about whether they would have been legally justified if the facts and circumstances had been as they believed indicates the necessity for narrow legalistic interpretation, which seems beside the moral point in these cases.

For the same reason, under Professor Slobogin's integrationist approach Andrea Yates should be convicted of negligent homicide for the deaths of her four children. Slobogin says that she "presumably would have been found to justify her criminal conduct,"[202] but this conclusion is simply incorrect as a matter of positive law, even if the facts and circumstances had been as Ms. Yates believed and if she were not responsible for causing the condition of her own excuse. Suppose Ms. Yates would have succeeded with the choice of evils justification. Should we excuse her because the law should hold that it is better children should die than be tormented in hell for eternity—in which case she would be excused under the "insane delusion" part of *M'Naghten* and no integrationist approach is required-- or because Ms. Yates was an irrational agent at the time of the homicide? Again, is it morally and legally sensible to be arguing what should be the right legal result if bizarre facts were true, or should we instead be focusing on the defendant's rational capacity as manifested by her abnormal, bizarre motivating beliefs? Yates may be a sympathetic defendant, but she would have a much better chance of being excused under a sensible interpretation of a cognitive insanity defense than under the integrationist approach. To be consistent, the integrationist approach could try to subjectivize the "choice of evils" justification, but this would permit defendants to define

---

[200] See, e.g., MPC, sec. 3.02(1)(a)
[201] MPC, sec. 3.02(2).
[202] MJ, at 25, 53-54.

the scope of their own excuse, a charge Professor Slobogin wrongly levels at rationality tests.

The integrationist approach would also excuse a defendant if duress would have obtained if the facts and circumstances had been as the defendant believed. Such cases are clinically rare, but the objection is similar to the objection to the integrationist "justification" standard. If the defendant is negligent in estimating whether he is in such a situation, the defense would be unavailable for crimes requiring negligence as the mens rea. Either the duress standard must be entirely subjectivized for both mentally normal and abnormal defendants—an undesirable legal outcome because it would acquit many culpable, mentally normal defendants--or the legal distinction between mentally normal and abnormal defendants must be preserved.

The final part of the integrationist proposal would excuse a defendant who acted under a "general ignorance of law," by which Slobogin means agents who may not be aware of the concept of crime or who cannot understand the concept of good and evil.[203] Virtually no normal defendant could sustain such a claim, and, among those with mental abnormalities, as a clinical matter, only psychopaths might be able to do so. Some deluded defendants may have entirely irrational views about the parameters of crime or of the concepts of good and evil, but few abnormal defendants other than psychopaths lack all understanding of these concepts Professor Slobogin criticizes irrationality tests for insanity as potentially excusing psychopaths,[204] but the integrationist approach will also excuse some psychopaths, as he himself recognizes.[205]

---

[203] MJ, at 27.

[204] MJ, at 42-44, 51.

Applied to psychopaths, this is not really an "integrationist" proposal at all because it affects only mentally abnormal defendants. It is in fact a form of legal insanity defense, but it is not the best approach to these cases. The best reason to excuse psychopaths is not a "general ignorance of law" defense. It is that psychopaths lack the capacities—empathy and guilt—that are the primary tools rational agents use to give them good reason not to harm others unjustifiably. They are irrational in that respect. If we wish to excuse psychopaths—and many do not--an interpretation of the insanity defense that permits this result is the most morally precise and common sense approach

Finally, there is a fundamental inconsistency in Professor Slobogin's "integrationist" approach. He applauds the Supreme Court decisions categorically barring execution of capital murderers who were retarded or 16 and 17 years old at the time of the crime,[206] and argues that it would be unconstitutional to execute defendants who suffered from mental disorder at the time of the crime. He also argues that mental abnormality should be a mitigating factor in sentencing generally. His primary rationale for these positions is that mental abnormality diminishes the defendant's rationality, and that such abnormality- based diminished rationality can be reasonably assessed. Just so. Slobogin thus concedes that irrationality is the primary mitigating and excusing condition. But of course lack of rational capacity is a continuum. If a substantial lack of rational capacity mitigates responsibility for purposes of just punishment, Professor Slobogin provides no reason to deny that extreme irrationality should excuse altogether[207]

---

[205] MJ, at 82 (making this point in a later chapter addressing the death penalty). Professor Slobogin claims to have recognized this in his earlier chapter on the insanity defense, but the citation is to an article by a different scholar and all discussion of psychopathy in the insanity defense chapter strongly implies that psychopaths should not be excused.

[206] Atkins v. Virginia, 536 U.S. 834 (2002); Roper v. Simmons, 543 U.S. 551 (2005). MJ, at 64-83.

[207] MJ, at 75-77,

A suitably drafted and interpreted insanity defense addresses the fundamental excusing condition of irrationality and provides more just results than the integrationist call for legalistic and occasionally bizarre decision-making criteria.

VI. The Future of Mens Rea and Legal Insanity

We hope we have demonstrated the critical importance, if not necessarily the constitutional pedigree, of both mens rea and legal insanity to any morally coherent system of criminal justice, and the unworkability of alternatives. But two problems remain. First, there are those—both in law and science—who claim that the moral premises of criminal responsibility are quaint and unscientific remnants of a bygone era, and that advances in neuroscience are dissolving notions of blameworthiness into meaninglessness.[208] Second, even if we successfully resist this materialist and determinist onslaught, where does that leave moralists? What are the foundations of a morally-grounded theory of criminal responsibility?

A. The General Challenge to Responsibility

Scientific, physicalist theories of the mind threaten our deeply-held presumptions of intentionality and personhood and thus threaten our responsibility concepts and practices. What is happening when we "act"? What is happening when we "intend" to act? What is happening when we subjectively feel and report that we are apparently unable to "control" our acts? Until relatively recently, these questions largely inhabited the halls of philosophy and speculation. These problems were technically daunting

because accessing brain structure and processes was generally rudimentary or impossible. Advances in technology now allow brain imaging that can measure the electrical and heat-producing activities of single neurons during an individual's designated activity, and the accumulated mapping of that neural activity over time and over many subjects. As a result, neuroscientists announce almost daily that certain activities seem to be associated with certain areas in the brain, and that new levels of functional interconnectedness between different parts of the brain have been discovered.[209]

Some people, including many neuroscientists, think that these discoveries are leading inexorably to a mechanistic, deterministic view of the brain, and thus to a mechanistic, deterministic view of human behavior. Daniel M'Naghten didn't shoot that gun, it was his finger that pulled the trigger, and his peripheral neurons that caused his finger muscles to contract, and central neurons that caused the peripheral neurons to fire, and neurons and neurons all the way down. M'Naghten's intentions did no work in explaining his action. This is the thought that terrifies many thinking people about scientific advances in the understanding of human behavior.

The seriousness of neuroscience's potential challenge to the traditional foundations of law and morality is best summed up in the title of an eminent psychologist's recent book, The Illusion of Conscious Will .[210] Its thesis, stated most

---

[208] E.g., Joshua Greene & Jonathan Cohen, *For the law, neuroscience changes nothing and everything,* in Semir Zeki & Oliver Goodenough, Eds., LAW & THE BRAIN 207 (2006); Robert Sapolsky, *The frontal cortex and the criminal justice system*, in Zeki & Goodenough, supra at 227.

[209] See, e.g., B. Knutson et al, *Neural predictors of purchases*, 53 NEURON 147 (2007) (the questions shoppers ask themselves, how much they want an object and how much they are willing to pay for it, are associated with activity in specific parts of the brain). As is so often the case, the media quickly and breathlessly pick up the story, usually concluding or implying that the agent is therefore not responsible for the activity in question. See e.g., John Tierney, *The Voices In My Head Say `But It!' Why Argue?* NY TIMES (January 16, 2007).

[210] Daniel Wegner, THE ILLUSION OF CONSCIOUS WILL (2002).

extremely, is that we delude ourselves when we think that our intentions are genuinely causal. The rich explanatory apparatus of intentionality is simply a post-hoc rationalization we hapless homo sapiens construct to explain what our brains have already done. We are just mechanisms, although the illusion of conscious will may play a positive role in our lives.[211] The evidence for these assertions is not all based on neuroscientific findings, and the claim that we are purely mechanisms has long been thought to follow from all physicalist, naturalist, scientific views of the person, but neuroscience is the source of the most profound challenge today.

If this particular materialist and determinist view of humans were right, that is, if we are really not acting agents at all, the present foundations of law and morality rooted in agentic personhood would collapse. If scientific investigation can demonstrate that the alleged criteria for personhood cannot be satisfied by people as they are, the criminal law's central tenet that individuals are personally responsible for their behavior would

---

[211] Here is an extensive quotation from the conclusion of Professor Wegner's book, which is necessary to obtain the tenor of the assertion and to evaluate if it is internally logical:

> Sometimes how things seem is more important than what they are. This is true in theater, in art, in used car sales, in economics, and–it now turns out–in the scientific analysis of conscious will. The fact is, it seems to each of that we have a conscious will. It seems we have selves. It seems we have minds. It seems we are agents. It seems we cause what we do. Although it is sobering and ultimately accurate to call all this an illusion, it is a mistake to conclude that the illusory is trivial. On the contrary, the illusions piled atop apparent mental causation are the building blocks of human psychology and social life. It is only with the feeling of conscious will that we can begin to solve the problem of knowing who we are as individuals, of discerning what we can and cannot do, and of judging ourselves morally right or wrong for what we have done.

Id. at 341-42.

Error! Main Document Only. In more recent work, Professor Wegner appears to have softened the radical interpretation of his claim, which is that we as persons, as agents, are not really "controllers" whose mental processes cause action. Daniel M. Wegner, *Who is the Controller of Controlled Processes?,* in **Error! Main Document Only.**Ran R. Hassin, James S. Uleman & John A. Bargh, Eds., THE NEW UNCONSCIOUS 19, 32 ("The theory is mute on whether thought does cause action.") On the other hand, Professor Wegner seems ambivalent and loathe fully to give up the radical interpretation. See, id. at 27 (arguing that the experience of conscious will is "normally a construction" and referring to mental causation as "apparent.")

disappear. But don't fret. This collapse—which materialists have been predicting for centuries—is not imminent. Advances in neuroscience and related fields have revealed hitherto unimagined biological causes that predispose people to behave as they do,[212] but the science typically supporting the more general and radical claim that conscious will is an illusion–that we do not act and therefore cannot be responsible--is either insufficient empirically or does not have the implications claimed.[213]

To begin, it is important to recognize that the new challenge to conscious intentionality is distinguishable from the more familiar challenge to responsibility that determinism or causation generally presents. Even if determinism is true, we may well be the types of creatures we take ourselves to be—usually conscious, potentially rational creatures whose behavior is explained by our intentions. Assuming that the truth of determinism is logically inconsistent with responsibility *tout court*--a controversial position in the philosophy of responsibility--we can never discover if determinism is true. Moreover, there is a respectable core position in the philosophy of responsibility—

---

[212] Recently published issues of prestigious journals contain illustrative examples of advances in scientific understanding of the causes of legally relevant behavior. See, e.g., Avshalom Caspi et al, *Role of Genotype in the Cycle of Violence in Maltreated Children*, 297 SCIENCE 851 (2002) (maltreated male children were more likely to exhibit antisocial behavior if they had a defect in the genotype that confers high levels of the neurotransmitter encoding enzyme, monoamine oxidase A, which metabolizes various neurotransmitters linked to violence if the levels of those neurotransmitters are low); **Error! Main Document Only.**Rita Z. Goldstein & Nora D. Volkow, *Drug Addiction and Its Underlying Neurobiological Basis: Neuroimaging Evidence for the Involvement of the Frontal Cortex*, 159 AM. J. PSYCHIATRY 1542 (2002) (addiction involves cortical processes that result in the overvaluation of drug reinforcers, the undervaluation of other reinforcers, and defective inhibitory control of responses to drugs); J. Kim-Cohen et al, MAOA, *maltreatment, and gene-environment interaction prediction children's mental health: new evidence and a meta-analysis*, 11 MOLECULAR PSYCHIATRY 903 (2006)(MAOA gene influences vulnerability to environmental stress, such as childhood maltreatment); Marc N. Potenza et al, *Gambling Urges in Pathological Gambling: A Functional Magnetic Resonance Imaging Study*, 60 ARCHIVES OF GENERAL PSYCHIATRY 828 (2003) (when viewing gambling cues, male pathological gamblers "demonstrate relatively decreased activity in brain regions implicated in impulse regulation compared with controls"); Murray B. Stein et al, *Genetic and Environmental Influences on Trauma Exposure and Posttraumatic Stress Disorder Symptoms: A Twin Study*, 159 AM. J. PSYCHIATRY 1675 (2002) (concluding that genetic factors can influence the risk of exposure to assaultive trauma and to PTSD symptoms that may ensue).

"compatibilism"—that holds that determinism is *not* inconsistent with responsibility.[214] The legal system is in any event unlikely to abandon responsibility altoghether based on a speculative metaphysical thesis. We therefore do not discuss the radical and general deterministic challenge to responsibility.

It is important to recognize, however, that simply discovering a cause for behavior, whether it is biological, psychological or sociological, does not mean that the agent is not responsible for the behavior. All behavior has causes. If causation were an excuse, no one could ever be held responsible for any behavior. One of us has called the confused, pernicious belief that causation excuses, which is used selectively and all too frequently, the "fundamental psycholegal error."[215] This error is conceptually identical to the misguided belief that determinism excuses in our positive law. On the other hand, science may in fact be able to demonstrate that we are not the types of intentional agents that we think we are and that alone among the phenomena of the universe are justifiably held responsible. If so, responsibility concepts would have to change radically. If these doubts are accurate, compatibilism cannot save responsibility because responsibility depends on agency and the new discoveries arguably deny the possibility of agency as it is traditionally conceived. We therefore limit ourselves to this form of a general radical challenge to responsibility because it is at least demonstrable in principle.[216]

Materialism's relentless gnawing at moralism is neither new nor restricted to science. The philosophy of mind and action has long contained arguments for various

---

[213] See infra.
[214] Stephen J. Morse, *Reason, Results and Criminal Responsibility,* 2004 ILL. L. REV. 363, 437ff.
[215]Stephen J. Morse, *Culpability and Control*, 142 U. PA. L. REV. 1587, 1592-94 (1994).
[216] See generally, Stephen J. Morse, *The Non-Free Will Problem in Forensic Psychiatry and Psychology*, note ** supra, (addressing the two general challenges to responsibility and concluding neither poses a threat at present).

forms of material reductionism and for eliminative materialism.[217] Both types of conceptual views existed long before recent discoveries in neuroscience and psychology deepened our understanding of how the brain and nervous system are constructed and work. Both types of views are extremely controversial.

Most philosophers of mind believe that complete reduction of mind to biophysical explanation is impossible.[218] Again, until the conceptual revolution that allows us to solve the mind-body problem occurs—and, at present, we have no idea how the brain enables the mind[219]--science cannot resolve the debate, although it can furnish support for various conceptual arguments. At present, and for the foreseeable future, we have no convincing conceptual reason from the philosophy of mind, even when it is completely informed by the most recent neuroscience, to abandon our view of ourselves as creatures with causally efficacious mental states.[220]

Even if we cannot solve the mind-body problem, however, and thus determine if reductive accounts are true, it is possible that we might make empirical discoveries

---

[217] Reductive accounts hold, simply, that mental states are as they seem to us, but that they are identical to brain states. Eliminative accounts hold that our beliefs about our mental states are radically false and, consequently, that no match between brain states and mental states is possible. See Paul M. Churchland, MATTER AND CONSCIOUSNESS 26-34, 43-49 (Rev. ed. 1988) (explaining the arguments for and against both types of accounts).

[218] *See*, e.g. John S. Searle, MIND: A BRIEF INTRODUCTION 88-92, 111-132 (2004); Galen Strawson, *Consciousness, Free Will, and the Unimportance of Determinism*, 32 INQUIRY 3 (1989) (claiming that reductive physicalism about the mind is "moonshine"); *see generally*, John R. Searle, THE REDISCOVERY OF THE MIND (1992) (providing an extended argument for the irreducible reality of mind).

[219] Paul McHugh & Philip Slavney, THE PERSPECTIVES OF PSYCHIATRY (2D ED.) 11-12 (1998).

[220] To claim that mental states may play a causal role in behavior does not commit one to discredited mind-brain dualism. Even if mental states can not be reduced to brain states at the level of neurons, the mind is produced by and realizable in the brain, a position that may be termed "non-reductive monism" or biological naturalism. See John Searle, note ** supra at 113-114. It is an almost inevitable feature of language to sound dualistic, however. For example, when critics of dualism talk about "my" or "your brain" they imply that there is a "me" or a "you," a self, that is distinguishable from one's brain. Imagine expressing that thought by talking about "the brain that is inside this skull [pointing to one's own head or the head of another]."

indicating that some parts of our ordinary understanding about action and agency are incorrect. Much recent argument based on current neuroscience and psychology takes this position, arguing that mental causation does not exist as we think it does. For ease of exposition, we label these approaches the "no action thesis" ("NAT"), as a way of describing their ultimate conclusion that the notion of human "action" is a neurological artifact of consciousness. But the logic of these arguments is often shaky. Discovering a brain correlate, or any other cause of an action, does not mean that it is not an action. If actions exist, they have causes, including those arising in the brain.

The real NAT question is whether scientific experiments have shown that action is rare or non-existent, that conscious will is largely or entirely an illusion after all. Three kinds of evidence are often adduced: First, demonstrations that a very large part of our activity is undeniably caused by variables of which we are not at all aware; and, second, studies indicating that more activity than we think takes place when our consciousness is divided or diminished, and third, laboratory studies that show that people can be experimentally misled about their causal contribution to their apparent behavior . Neither kind of evidence offers logical support to NAT, however.

Just because a person may not be aware of all the causes for why he formed an intention does not mean that he did not form an intention and was not a fully conscious agent when he did so. Even if human beings were never aware of the causes of their intentions to act and actions, it would not necessarily follow that they were not acting consciously, intentionally and for reasons that make eminent sense to anyone under the circumstances.

Human consciousness can undeniably be divided or diminished by a wide variety of normal and abnormal causes.[221] We have known this long before contemporary scientific discoveries of what causes such states and how they correlate with brain structure and processes. Law and morality agree that if an agent's capacity for consciousness is non-culpably diminished, responsibility is likewise diminished. As we have seen,[222] some suggest that it is diminished because bodily movements in the absence of fully integrated consciousness are not "actions." Others believe that apparently goal-directed behavior that is responsive to the environment, such as sleepwalking, is action, but that it should be excused because diminished consciousness reduces the capacity for rationality. Let us assume that the former view is correct, because it offers more direct support to NAT, and therefore the greatest challenge to traditional notions of individual responsibility. Let us also assume that divided or diminished consciousness is more common than we think. Nevertheless, neither of these assumptions supports the more radical, general NAT thesis.

To demonstrate that divided or partial consciousness is more common than it appears certainly extends the range of cases in which people are not responsible or have diminished responsibility, but such studies do not demonstrate that most human bodily movements that appear intentional and rational occur when the person has altered consciousness. One cannot generalize to all human behavior from genuinely deviant cases or cases in which a known abnormality is present. A model of action (or, we

---

[221]See, Jeffrey L. Cummings & Michael S. Mega, NEUROSPSYCHIATRY AND BEHAVIORAL NEUROSCIENCE 333-43 (2003) (description of dissociative and related states and their causes and treatments). D. Vaitl , N. Birbaumer, J. Gruzelier *et al* , *Psychobiology of altered states of consciousness*, 131 PSYCHOLOGICAL BULLETIN 98 (2005).

[222] Note *** supra and accompanying test.

should say, non-action) built on sleepwalking, for example, is hardly a threat to orthodox notions of individual responsibility.

Last, there is substantial empirical evidence to suggest that experimental laboratory manipulations of unsuspecting subjects can cause the subjects to believe that their intentions were producing action when this was not the case.[223] That subjects can be cleverly mislead to misunderstand the causal role of their intentionality by experimental manipulations hardly indicates that intentions generally play no role in explaining our behavior. The brain and other material variables may cause our behavior, but not because such variables are intentional manipulators.[224]

What is needed to support NAT is a general demonstration that causal intentionality is an illusion *tout court*, but no such general demonstration has yet been produced by scientific study. The most interesting evidence has arisen from studies done by Benjamin Libet,[225] which have generated an immense amount of comment.[226] Libet's exceptionally creative and careful studies demonstrate that measurable electrical brain activity associated with intentional actions occurs about 550 milliseconds before the subject actually acts and for about 350-400 milliseconds before the subject is consciously aware of the intention to act.

---

[223] See, John A. Bargh, *Bypassing the Will: Toward Demystifying the Nonconscious Control of Social Behavior*, in Hassin et al, note *** supra at 37(reviewing the evidence and concluding that the "will" is not primarily responsible for action).

[224] Compare, Daniel C. Dennett, ELBOW ROOM: THE VARIETIES OF FREE WILL WORTH WANTING 8-9 (1984)(explaining in the related context of the determinist critique that determinism does not mean that a "nefarious neurosurgeon" is controlling your behavior).

[225] Benjamin Libet, *Do We Have Free Will*, in Benjamin Libet, Anthony Freeman & Keith Sutherland, eds., THE VOLITIONAL BRAIN: TOWARDS A NEUROSCIENCE OF FREE WILL 47 (1999) (summarizing the findings and speculating about their implications).

[226] Wegner, supra note___ at 54-55 (characterizing the recounting of Libet's results as a "cottage industry" and noting the large and contentious body of commentary).

Let us assume, with cautious reservations,[227] the validity of these studies. It does not follow from this temporal ordering that conscious intentionality does no causal work. It simply demonstrates that non-conscious brain events precede conscious experience, but this seems precisely what one would expect of the mind-brain. It does not mean that intentionality played no causal role, and Libet concedes that people can "veto" the act, which is another form of mental act that plays a causal role. Finally, Libet's task involved "random" finger movements that involved no deliberation whatsoever and no rational motivation. This is a far cry from the type of intentional conduct that is the criminal law's concern. Libet's work is fascinating, but it does not provide a general proof that humans are never or infrequently conscious, intentional agents.[228]

Moreover, NAT provides no guidance about what we should do next, and in any event degenerates into a self-referential incoherence. Let us suppose that you were convinced by the mechanistic view that you were not an intentional, rational agent after all. (Of course, the notion of being convinced would be an illusion, too.) What should you do now? You know it is an illusion to think that your deliberation and intention has any causal efficacy in the world. You also know, however, that you experience sensations such as pleasure and pain and that you care about what happens to you and to the world. You cannot just sit quietly and wait for your neurotransmitters to fire. You must and will of course deliberate and act.

That is, even if pure mechanism is true–about which, once again, we will never be certain until we solve the mind-body problem–human beings will find it almost

---

[227] E.g., Henrik Walter, NEUROPHILOSOPHY OF FREE WILL: FROM LIBERTARIAN ILLUSIONS TO A CONCEPT OF NEURAL ANATOMY (Cynthia Klor tr.).250-252 (2001); J. Zhu, *Reclaiming volition: an alternative interpretation of Libet's experiment*, 10 JOURNAL OF CONSCIOUSNESS STUDIES 61 (2003).

impossible not to treat themselves as rational, intentional agents unless there are major changes in the way our brains work. Indeed, as discuss below, there are powerful evolutionary explanations for the causal efficacy of our mental states, states that are the predicates for responsibility.[229] Overcoming what may in fact be neuro-architectural facts of evolution may be only a little less likely than "overcoming" our opposable thumbs.[230] Moreover, as we saw, if you use the truth of pure mechanism as a premise in deciding what to do, this premise will entail no particular moral, legal or political conclusions. It will provide no guide to how one should live, including how one should respond to the truth of NAT.

Finally, the argument from common sense in favor of the justified belief that we are conscious, intentional creatures is overwhelming. Consider again, for example, the nature of law itself. As we have seen, law is a system of rules that at the least is meant to guide or influence behavior and thus to operate as a potential cause of behavior. It would be impossible at present for us to abandon the well-justified belief that action may be influenced by reason and that our intentions are causally efficacious. As the eminent philosopher of mind, Jerry Fodor wrote:,

> [I]f commonsense intentional psychology were really 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, doesn't compare…Nothing except, perhaps, our commonsense physics… comes as near our cognitive core as intentional explanation does. We'll be in deep, deep trouble if we have to give it up… But be of good cheer; everything is going to be all right.[231]

---

[228] See also, T. Crane, *Ready or not*. 5311 TIMES LITERARY SUPPLEMENT 26 (January 14, 2005).

[229] See infra.

[230] See text accompanying notes ** to ** infra.

The new neuroscience does not pose, and is unlikely ever to pose, a real threat to our fundamental conception of personhood and all that follows from it, including the concept of responsibility and related concepts, such as mens rea. At the very least, we remain entitled to presume that conscious intentions are potentially rational and causal, and to place the burden of persuasion at a very high level on proponents of NAT. At present, the case is not close to meeting the burden. We are fully entitled to continue to believe that we are the sorts of creatures we believe ourselves to be.

B. The Roots of Responsibility

If the sciences have not yet shown that individual responsibility is an unjustified illusion, where does it come from? And might those roots provide a useful guide to what the criteria should be for holding each other responsible? In this section we consider the newest philosophical and neuroscientific approaches to this age-old question.

One of the most influential modern philosophical theories of responsibility, beginning with a classic article by P.F. Strawson[232] and expanded by many others,[233] argues that responsibility is rooted in our reactive emotions, such as indignation,

---

[231] Jerry Fodor, PSYCHOSEMANTICS: THE PROBLEM OF MEANING IN THE PHILOSOPHY OF MIND xii (1987). The entire book is a defense of commonsense intentional explanation. See generally, Bertram F. Malle, HOW THE MIND EXPLAINS BEHAVIOR: FOLK EXPLANATIONS, MEANING AND SOCIAL INTERACTION (providing a full theoretical account and empirical support). There is also growing recognition within psychology that "mental state inference is one of the most fundamental tools of social cognition." Bertram F. Malle, *Folk Theory of Mind: Conceptual Foundations of Human Cognition,* in Ran R. Hassin et al, note ** supra, at 225, 229

[232] *Freedom and Resentment*, in Gary Watson, Ed., FREE WILL 59 (1982).

[233] E.g., R. Jay Wallace, RESPONSIBILITY AND THE MORAL SENTIMENTS (1994).

resentment, anger, and gratitude. When we are benefited or harmed by compliance with or breach of a normative obligation, we tend to have such reactive emotions, but we think that such reactions are justified only if the agent who helped or hurt us acted with the right mental state and was responsible. Further, if we think that our reactive attitude is justified, we also think that we are justified in expressing that attitude in an appropriate way. This process can occur publicly or privately, as personal moral criticism or criminal law blame and punishment, for example. To hold someone responsible, then, is to believe that one is justified in having and expressing a reactive attitude. In criminal responsibility terms, holding responsible means that we believe we are justified in blaming and punishing wrongdoers. Blaming in this context means more than simply grading. It refers to a moral evaluation of the criminal's conduct.[234]

Although the content of the normative obligations and the specific reactive emotions and expressions might vary temporally or culturally, it is probably true that all cultures have responsibility practices that bear some resemblance to this account. It is certainly the case in western systems of morality and law. It is also striking that there is enormous agreement, including cross-culturally, about the rank ordering of the seriousness of the harms to persons and property that are at the core of the criminal law,[235] although much less agreement about the appropriate penalty structure.

Is there an even deeper explanation of the ubiquity of the reactive emotions themselves? Yes, and not without a great deal of irony, those deeper explanations are coming from the sciences—from evolutionary biology, evolutionary psychology and

[234] Bok, supra note___ at ____.

indeed even from neuroscience itself. In fact, Strawson predicted some of these results, asserting that the reactive attitudes were part of human nature, and that it would be impossible to obliterate them, and undesirable to do so even if we somehow could.[236] Having biological roots of course does not guarantee that a practice is normatively desirable.[237] To believe otherwise is the naturalistic fallacy. But if a practice has biological roots, biology can indicate the "price" for abandoning it.

Evolutionary theorists have demonstrated with game-theoretic precision that many otherwise confounding animal behaviors, including some human behaviors, are perfectly adaptive; that is, they are widespread because they provided individual actors with survival and thus reproductive advantage.[238] It is now settled that all animals, including humans, have evolved tendencies to act in certain predictable ways under certain conditions.

We believe that this sort of evolutionary account is the best explanation of the ubiquity of the reactive responsibility practices and of the powerful agreement about ordinal harm rankings. There is good experimental evidence to suggest, for example, that homo sapiens are predisposed to be altruistic retributivists.[239] A more recent

---

[235] Paul H. Robinson & Robert Kurzban, *Concordance & Conflict in Intuitions of Justice* (ms. 2006) (peoples' intuitions of justice concerning the core harms of the criminal law are specific, nuanced and widely shared cross-culturally).

[236] See also, P.F. Strawson, SKEPTICISM AND NATURALISM: SOME VARIETIES 35-37 (1983). Later commentators have differed about a seemingly naturalized account of the reactive emotions account of responsibility. Compare George Sher, IN PRAISE OF BLAME 1 (2005) (criticizing Strawson) with Andrew Oldenquist, *An Explanation of Retribution*, __ J. PHIL. ___ (1988) (arguing for a naturalized account of retribution).

[237] See Bok, note *** supra at 26-27.

[238] See, e.g., W.D. Hamilton, The Genetical Evolution of Social Behavior, 7 J. THEOR. BIOL. 1-52 (1964); Robert L. Trivers, The Evolution of Reciprocal Altruism, 46 Q. REV. BIOL. 35-37 (1971).

experimental study demonstrates that societies with an altruistic punishment mechanism have a competitive advantage compared with a rival society without one. Although subjects initially prefer the society without punishment, they quickly migrate to the sanctioning society because greater social cooperation occurs in the latter.[240] All these findings are consistent with an evolutionary account of how successful human societies enforced cooperation by their response to free riding.[241]

But punishing free-riders is just half the evolutionary story. We have adaptive proclivities to free-ride, but we also have adaptive proclivities to cooperate. As a profoundly social species, it takes no great leap of imagination to understand that our ancestors would have been much less successful, and therefore much less likely to be our ancestors, if they were not sure they could convert their intentions into action, and were not also equipped with powerful assumptions that that stranger coming over the hill is equipped with the same capacity for harmful intentions as they are. Moreover, evolution has equipped us with an advanced and sensitive system to modulate short term individual benefits (steal that car) against long term individual costs (go to prison), and vice versa. They are two sides of the same evolutionary coin, and, inflamed with that sometimes

---

[239] *See* Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137, 137 (2002) (experimental evidence suggests that human beings have evolved to engage in nonutilitarian, "altruistic" punishment when members of their social group "defect" by violating a normative expectation). More recent, consistent evidence suggests that, at least in men (who do most of the punishing for free riding), empathic responses are shaped positively by fair treatment by others and negatively by unfair treatment. Tania Singer et al, *Empathic neural responses are modulated by the perceived fairness of others,* NATURE (pub. online 18 Janurary 2006). Other cooperative primates exhibit behavior best interpreted as expressing resentment when they perceive that they are treated undeservedly unequally. *See* Sarah F. Brosnan & Frans B. M. de Waal, *Monkeys Reject Unequal Pay*, 425 NATURE 297, 297 (2003). See generally, Paul W. Glimcher, DECISIONS, UNCERTAINTY AND THE BRAIN: THE SCIENCE OF NEUROECONOMICS (MIT 2003).

[240] Ozgur Gurerk, Bernd Irlenbusch & Bettina Rockenbach, *The Competitive Advantage of Sanctioning Institutions*, 312 SCIENCE 108 (2006)

[241] David M. Buss, EVOLUTIONARY PSYCHOLOGY: THE NEW SCIENCE OF THE MIND (2D ED.) 276-77 (2004). Indeed, that irrepressible sense of free will and "personhood" that we have already discussed in philosophical and legal contexts is most likely irrepressible because our brains evolved to experience and

cursed and almost-ever-present consciousness, they are the heart of the human dilemma, and at the heart of law and morality.

Of course, holding people responsible and giving them their just deserts is clearly a complex social practice that is inevitably and powerfully affected by culture and individual learning and socialization. Blaming and punishing are not solely the product of our DNA or our neuro-architecture. There is no single punishment gene or neural network. Nonetheless, we seem strongly predisposed to feel and to express retributive sentiments, a predisposition that might not be so easy to overcome, even it were normatively a good idea to do so.

As we have suggested previously, holding people responsible and excusing others is profoundly normatively desirable. It takes people seriously as potentially moral agents and contributes to human dignity and respect. It also facilitates social cooperation. We would be giving up much that was worthwhile if we abandoned the practice of responsibility merely to be consistent with an allegedly clear-eyed and scientific view that no one should be blamed and punished and that we should be concerned solely with the consequences of social control. An irrationality excuse, like the insanity defense, may ultimately not be of much benefit in terms of social safety, but it is crucial to taking responsibility seriously.

VII. Mens Rea, Legal Insanity and Common Sense

The moral (pardon the expression) of our story is that the criminal law must continue to emphasize the importance of mens rea and rationality, and that deciding

---

construct that sense of free will and personhood. See, e.g., Morris B. Hoffman & Timothy H. Goldsmith, *The Biological Roots of Punishment*, 1 OHIO ST. J. CRIM. L. 627 (2004).

whether a defendant lacked mens rea or was sufficiently irrational to be excused are factual and moral decisions that may safely be left to the common sense of ordinary people.

There will still be interpretive questions, of course. Consider the following cases of potential legal insanity:

> 1) The delusional self-defender who genuinely believes that he is in imminent wrongful danger of death;
>
> 2) Daniel M'Naghten, who is equally deluded about the danger to his life, but does not believe that his death is imminently threatened by the victim;
>
> 3) A person who hallucinates that she had heard God's voice command her to kill her children and delusionally believes that she must obey in order to produce eternal peace on earth;
>
> 4) Andrea Yates, who believes that she must kill her children now to save them from Satan's eternal torments
>
> 5) A person who delusionally believes his spouse is unfaithful and kills her in a state of jealous rage sparked by his delusional belief.
>
> 6) The psychopath: a person who knows what he is doing and knows the "prices," such as going to prison, for failing to comply with certain rules, but who entirely lacks conscience and empathy and does not get the point of morality or its rules at all.

The first five defendants are equally deluded and equally irrational, but of course all meet the elements of the prima facie case, since all are at least arguably capable of forming the required mens rea. The psychopath clearly has no problem forming mens rea. So only legal insanity is in question. Yet there is no obvious right answer to these cases under any insanity defense rule. Indeed, the authors of this paper disagree about some of them.

If the question is simply whether the practical reason resulting in the crime was substantially affected by irrationality, then the first five should all be acquitted by reason

of insanity. If, instead, the test is whether the defendant acted for a good justifying reason under positive law if the deluded reasons were true,[242] then only some would be acquitted. Even then, some of these cases are close, including 3) and 4). In the latter, would the general justification of necessity or "balance of evils" apply if the belief were true? Is this even a sensible way to put the question? Finally, disposition of the sixth case depends entirely on how one defines rationality for purposes of ascribing responsibility. These are people who know what they are doing and know the rules, but who simply do not understand the difference between good and evil. Are they irrational or simply bad?

In the end, reasonably crafted definitions of mens rea and insanity, coupled with the common sense of jurors, should be sufficient to create systems of responsibility robust enough to reflect a strong presumption of blameworthiness, yet sensitive enough to recognize the rare exception. Indeed, jurors alone, with available elemental and excusing doctrines but with no other guidance about the details of those doctrines, seem perfectly competent to make such judgments.[243]

There is good evidence to suggest that the language of the insanity defense does not make a great deal of difference to the outcome of insanity cases.[244] Jurors apparently reach a rather "rough and ready" judgment about which offenders deserve to be excused. Not surprisingly, the major variable that affects the outcome is the severity of the

---

[242] This alternative is quite akin to the "justification" part of the integrationist approach discussed at Note ** supra and accompanying text, but in this case the excuse is explicitly tied to the agent's irrationality and no general subjectivization of justifications is necessary.

[243] Rita J. Simon & David E. Aaronson, THE INSANITY DEFENSE: A CRITICAL ASSESSMENT OF LAW AND POLICY IN THE POST-HINCKLEY ERA 113-135 (1988)(jurors could fairly and sensibly apply the law).

[244] Henry M. Steadman et al, BEFORE AND AFTER HINCKLEY: EVALUATING INSANITY DEFENSE REFORM __ (1993).

offender's disorder, that is, how grossly out of touch with reality the defendant was.[245] Non-psychotic offenders rarely succeed, which is the morally correct and intuitive result, since, as we have discussed, most non-psychotic people are not sufficiently irrational to deserve an excuse.[246]

In sum, we can apparently trust the good judgment of juries in most insanity cases. We cannot and should not expect more.

VIII. Conclusion

Clark reached the right result on legal insanity but the wrong result on mens rea, and in doing so perpetuated the blurring of the two. Stout, but clearly separate, doctrines of intentionality and excuse will help jurors use their common sense about both. That will in turn help insure results that mesh with our most fundamental notions of individual responsibility and blameworthiness. And, for now, desirable concepts of responsibility and blame are not undermined by the newest scientific findings about behavior.

---

[245] Steadman et al, supra note___ at ___.

[246] There is great behavioral heterogeneity among people with psychotic disorders, so not all of them should be excused, even if their psychotic processes played a role in their practical reason. DSM-IV-TR, note *** supra, at xxix.