Maryland Law Review

Volume 82 | Issue 3 Article 5

Guilty Minds

Michael Serota

GUILTY MINDS

MICHAEL SEROTA*

This Article develops a new vision of mens rea by returning to a bygone era's conception of the guilty mind. The common law understanding of mens rea is broad and moralistic; it encompasses all mental characteristics bearing on an actor's blameworthiness. Undertheorized and oft neglected, this "Guilty Minds" approach has been replaced with the Model Penal Code's reconceptualization of mens rea as the purpose, knowledge, recklessness, or negligence applicable to every element of a criminal offense. Modern criminal law's embrace of this narrower and more legalistic "PKRN" approach to mens rea has brought with it many well-known benefits. But there are also overlooked costs of divorcing mens rea doctrine from its moral foundations. This Article demonstrates how the Guilty Minds approach, once clarified and refined, can address these costs while revealing a promising new pathway for criminal law reform.

Synthesizing a wide body of experimental research, the Article transforms the historically vague Guilty Minds approach into a multi-dimensional model of culpability rooted in the community's sense of justice. Drawing on contemporary criminal theory, the Article then makes the moral philosophical case for viewing this reconceptualization of mens rea as a critical constraint on criminal liability. After identifying structural flaws in contemporary mens rea policies that violate this constraint, the Article proposes a novel statutory solution: an insufficient blameworthiness defense, which empowers factfinders to dismiss charges based upon a structured assessment of an accused's mitigating mental states. The Article argues that the proposed defense would be accessible to juries, administrable by courts, and harmonious with the PKRN approach—thereby providing all U.S. jurisdictions with an effective way to bolster mens rea protections in their criminal codes.

* Associate Professor, Loyola Law School; Director, Criminal Justice Reform Lab; Senior Research Scholar, Academy for Justice. I am grateful to Rachel Barkow, Doug Berman, Jenny Carroll, Laura Coordes, Colin Doyle, Joshua Dressler, Brenner Fissell, Steve Garvey, Doug Husak, Zak Kramer, Erik Luna, Kaipo Matsumura, Ben McJunkin, Matt Mizel, Trevor Reed, Andrea Roth, Alex Sarch, Richard Schmechel, Josh Sellers, Bijal Shah, Ken Simons, Michelle Singer, Justin Weinstein-Tull, and Gideon Yaffe for helpful comments and insightful feedback. I am especially thankful to Carly Brody, Luca Artista, and the rest of the student editors at the *Maryland Law Review* for their diligent work on this Article.

INTRODUCTION

Mens rea is equal parts foundation and paradox. The basic idea behind mens rea—that we should only blame wrongdoers who act with a guilty mind—is widely considered to be one of the great tenets of law. For example, "[w]estern civilized nations have long looked to the wrongdoer's mind to determine both the propriety and the grading of punishment,"[1] while some have argued that the evolution of mens rea reflects humanity's "continuing

1. United States v. Cordoba-Hincapie, 825 F. Supp. 485, 489 (E.D.N.Y. 1993); *see, e.g.*, Francis Bowes Sayre, *Mens Rea*, 45 HARV. L. REV. 974, 974 (1932) ("For hundreds of years the books have repeated with unbroken cadence that *Actus non facit reum nisi mens sit rea*. 'There can be no crime . . . without an evil mind.'" (footnote omitted) (quoting 1 JOEL PRENTISS BISHOP, BISHOP ON CRIMINAL LAW § 287 (9th ed. 1923))); ELIZABETH PAPP KAMALI, FELONY AND THE GUILTY MIND IN MEDIEVAL ENGLAND 3–11 (2019).

process of self-civilization."[2] More than just a historical practice or sociological lens, however, mens rea is a policy choice touted by courts and legislatures in the strongest terms possible.[3] "[U]niversal and persistent in mature systems of law," mens rea is declared to be as fundamental to the criminal law as our belief in an individual's ability to "choose between good and evil"[4] and "essential if we are to retain 'the relation between criminal liability and moral culpability' on which criminal justice depends."[5]

While U.S. legal culture portrays mens rea as sacrosanct, the nation's criminal codes offer a contrasting vision. From strict liability offenses that authorize criminal liability in the absence of any culpable mental state,[6] to sentencing statutes that set the scale of punishment without regard to critical moral differences between culpable mental states,[7] neglect of the guilty mind in U.S. criminal legislation is pervasive. The result is a legal system in which the morally innocent are convicted for reasonable mistakes,[8] while the minimally blameworthy receive extreme sentences in light of unforeseen occurrences that transpire during the perpetration of crimes.[9] And these are just the most well-known ways that American criminal law neglects the

2. Paul H. Robinson, *A Brief History of Distinctions in Criminal Culpability*, 31 HASTINGS L.J. 815, 850 (1980). *See generally* MORRIS B. HOFFMAN, THE PUNISHER'S BRAIN: THE EVOLUTION OF JUDGE AND JURY (2014) (exploring the relationship between human evolution and moral intuitions relevant to mens rea assessments and policy choices).

3. *See* Michael Serota, *How Criminal Law Lost Its Mind*, BOS. REV. (Oct. 27, 2020), http://bostonreview.net/law-justice/michael-serota-how-criminal-law-lost-its-mind ("If legal rhetoric were an accurate gauge of legislative reality, our criminal justice system's treatment of mens rea would be pristine.").

4. Morissette v. United States, 342 U.S. 246, 250 (1952).

5. Tison v. Arizona, 481 U.S. 137, 171 (1987) (Brennan, J., dissenting) (quoting People v. Washington, 402 P.2d 130, 134 (Cal. 1965)).

6. *See, e.g.*, Michael Serota, *Strict Liability Abolition*, 98 N.Y.U. L. REV. 112 (2023) (highlighting the breadth of pure strict liability in U.S. criminal codes).

7. *See, e.g.*, Michael Serota, *Proportional Mens Rea and the Future of Criminal Code Reform*, 52 WAKE FOREST L. REV. 1201, 1202 (2017) (discussing the routine neglect of mens rea distinctions in legislative grading determinations).

8. *See, e.g.*, Catherine L. Carpenter, *On Statutory Rape, Strict Liability, and the Public Welfare Offense Model*, 53 AM. U. L. REV. 313 (2003) (highlighting the extent to which jurisdictions reject a reasonable mistake of age defense in statutory rape prosecutions); Laurie L. Levenson, *Good Faith Defenses: Reshaping Strict Liability Crimes*, 78 CORNELL L. REV. 401, 406–17 (1993) (discussing how people who make reasonable efforts to assess the age of employees can be convicted under federal law).

9. *See, e.g.*, RACHEL ELISE BARKOW, PRISONERS OF POLITICS: BREAKING THE CYCLE OF MASS INCARCERATION 25 (2019) ("Forty-six jurisdictions have laws that enable prosecutors to charge someone with murder if he or she commits a felony and someone [accidentally] dies during the course of that felony . . . ."); Michael Serota, *Second Looks & Criminal Legislation*, 17 OHIO ST. J. CRIM. L. 495, 509 n.77 (2020) ("[I]n many jurisdictions, a person who purposely assists with, or conspires in, the commission of one crime may be held fully responsible for any other reasonably foreseeable crimes committed by the principal actor—in the absence of the supporting actor's subjective culpability . . . .").

minds of wrongdoers. Many others are less recognizable to those who make criminal policy, yet are equally problematic.

For example, every year, untold masses of people are convicted of crimes without an accounting of why they did what they did, or whether their decisions were impaired by circumstances they could not control.[10] Motivated by good intentions, some of these individuals may have acted to further societally-legitimate objectives—for example, to avoid harm to oneself or others. Others charged with criminal wrongdoing may have been subjected to severe psychological stressors—from mental illness, to external coercion, to justified anger—that made it more difficult for them to think or act morally. The ethical salience of these mental states is easily recognizable to the public,[11] and their tendency to mitigate blame is well-established in our social morality.[12] But when these mitigating mental states arise in the context

10. This is made possible by the fact that the same psychological phenomena (e.g., motives, rationality, and volitional control) that provide the basis for complete defenses (e.g., insanity, duress, and self-defense) are usually ignored at the liability stage of criminal prosecutions whenever they fall short of satisfying a defense. This legal reality is further explored *infra* Sections I.B and IV.A. For previous scholarly discussions, see, for example, Serota, *supra* note 7, at 1215–16 (citing Douglas Husak, *Partial Defenses*, *in* THE PHILOSOPHY OF CRIMINAL LAW 311 (2010)) (discussing how the mitigating impact of partial defenses such as extreme mental or emotional disturbance is ignored outside the homicide context); Adam J. Kolber, *The Bumpiness of Criminal Law*, 67 ALA. L. REV. 855, 869 (2016) ("Criminal defenses . . . tend to be bumpy. If you satisfy the threshold of every element of the defense, you receive no punishment at all. But if you satisfy every element except one and are close to satisfying that element as well, you nevertheless lose the defense."); Avlana K. Eisenberg, *Discontinuities in Criminal Law*, 22 THEORETICAL INQUIRIES L. 137, 138 (2021) ("Bumpy relationships between inputs and outputs may seem particularly unfair when line drawing creates all-or-nothing results that fail to account for how close or distant a person's behavior was to the line in question."); Paul H. Robinson, *Mitigations: The Forgotten Side of the Proportionality Principle*, 57 HARV. J. ON LEGIS. 219 (2020).

11. For discussion of mens rea's role in community assessments of moral and criminal responsibility, see *infra* Part II.

12. For discussion of mens rea's philosophical foundations, see *infra* Part III.

of criminal prosecutions, they are typically ignored,[13] having been rendered legally immaterial by our criminal codes.[14]

What explains this gulf between the rhetoric and reality of U.S. mens rea policy? Undoubtedly, the same pathologies that wrought mass incarceration[15] have a role to play in understanding the criminal law's neglect

13. I say "typically" because there are exceptions. For example, the law of homicide recognizes some mitigating motives and cognitive impairments in making grading distinctions between murder and manslaughter. *See, e.g.*, Serota, *supra* note 7, at 1215–16. However, insofar as threshold determinations of criminal liability are concerned (the focus of this Article), the Model Penal Code's widely adopted definitions of recklessness and negligence incorporate a holistic blameworthiness analysis that accounts for mitigating mental states. MODEL PENAL CODE § 2.02(2)(c) & (d) (AM. L. INST. 1985) (defining recklessness and negligence); *see infra* Section IV.B.1 (discussing this holistic blameworthiness analysis). The Model Penal Code also recommends a *de minimis* provision empowering judges to dismiss cases, which some state courts have interpreted to entail consideration of mitigating mental states. MODEL PENAL CODE § 2.12 (AM. L. INST. 1985); *see infra* Section IV.B.2 (discussing relevant legal trends). Finally, mitigating mental states are sometimes recognized at sentencing, see Carissa Byrne Hessick & Douglas A. Berman, *Towards A Theory of Mitigation*, 96 B.U. L. REV. 161 (2016), albeit in ways that are often highly discretionary and procedurally suspect, see Serota, *supra* note 7, at 1218 (observing that American sentencing policies "leave trial courts with unchecked discretion over whether to give effect to mens rea-related aggravating or mitigating facts"); John B. Meixner Jr., *Modern Sentencing Mitigation*, 116 NW. U. L. REV. 1395, 1402 (2022) ("[C]urrent procedural structures [in the U.S.] do not encourage significant mitigation in most cases.").

14. As further discussed *infra* Section IV.A, this holds true even where multiple mitigating mental states exist—for example, committing a crime in the face of both significant mental illness and external coercion—provided none individually rises to the level of a complete defense.

15. The phrase "mass incarceration" can be construed in various ways. *See, e.g.*, Andrew D. Leipold, *Is Mass Incarceration Inevitable?*, 56 AM. CRIM. L. REV. 1579, 1581 (2019) ("The problem is that 'mass' incarceration does not have a fixed meaning."); Benjamin Levin, *The Consensus Myth in Criminal Justice Reform*, 117 MICH. L. REV. 259, 262 (2018) (arguing that despite the prevalence of the phrase mass incarceration "in scholarly and policy discussions, [it remains] ill-defined"). This Article's use of the phrase is intended to reference the following realities:

> Every morning in the United States 2.2 million people wake up in our nation's prisons and jails, making us the world leader in incarceration. Another 4.4 million people currently live under some form of correctional supervision, and more than 12 million cases flood our courts, public defender offices, and probation offices each year. The result is a country where one out of every three adults possesses a criminal record—a staggering statistic. Of all the lives, families, and communities destroyed by this system, the poor, the underserved, and people of color bear the brunt.

Serota, *supra* note 3; *see, e.g.*, Michael Serota, *Improving Criminal Justice Decisions*, 52 ARIZ. ST. L.J. 693, 695–97 (2020) (summarizing data on racial disparities in the criminal legal system) [hereinafter Serota, *Improving Criminal Justice Decisions*].

of the guilty mind.[16] Racist ideologies, beliefs and motivations;[17] the politics of law and order;[18] and the misguided notion that increased punitiveness invariably yields greater public safety[19] have all contributed to hostility toward mens rea.[20] But there is more to the story than just socio-political dynamics and the one-way ratchet that has become U.S. criminal law. Fundamental misconceptions matter too.

The topic of mens rea is subject to widespread confusion, even over basic issues. For example, which mental states count as mens rea under settled doctrine? There is no consensus answer. Although case law and scholarship on the topic is voluminous, courts and scholars perpetually struggle to identify "what the mens rea is."[21]

Confusion similarly surrounds the public's views on the topic. For example, what constitutes psychological blameworthiness in the eyes of the

16. For excellent overviews of these pathologies, see generally BARKOW, *supra* note 9; William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505 (2001); Sara Sun Beale, *What's Law Got to Do with It? The Political, Social, Psychological and Other Non-Legal Factors Influencing the Development of (Federal) Criminal Law*, 1 BUFF. CRIM. L. REV. 23 (1997). For an insightful recent discussion of how criminal law localism functions as a one-way ratchet for increasing criminalization, see Brenner M. Fissell, *Against Criminal Law Localism*, 81 MD. L. REV. 1119, 1138–46 (2022).

17. *See generally, e.g.*, Devon W. Carbado, *Blue-on-Black Violence: A Provisional Model of Some of the Causes*, 104 GEO. L.J. 1479 (2016); Kimberlé W. Crenshaw, *From Private Violence to Mass Incarceration: Thinking Intersectionally About Women, Race, and Social Control*, 59 UCLA L. REV. 1418 (2012); MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS (2010); Vesla M. Weaver, *Frontlash: Race and the Development of Punitive Crime Policy*, 21 STUD. AM. POL. DEV. 230 (2007).

18. *See generally, e.g.*, MARIE GOTTSCHALK, CAUGHT: THE PRISON STATE AND THE LOCKDOWN OF AMERICAN POLITICS (2015); WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE (2011); VANESSA BARKER, THE POLITICS OF IMPRISONMENT: HOW THE DEMOCRATIC PROCESS SHAPES THE WAY AMERICA PUNISHES OFFENDERS (2009); DAVID GARLAND, THE CULTURE OF CONTROL: CRIME AND SOCIAL ORDER IN CONTEMPORARY SOCIETY (2001).

19. *See generally, e.g.*, JONATHAN SIMON, GOVERNING THROUGH CRIME: HOW THE WAR ON CRIME TRANSFORMED AMERICAN DEMOCRACY AND CREATED A CULTURE OF FEAR (2007); Rachel E. Barkow, *Categorical Mistakes: The Flawed Framework of the Armed Career Criminal Act and Mandatory Minimum Sentencing*, 133 HARV. L. REV. 200 (2019); NAT'L RSCH. COUNCIL, THE GROWTH OF INCARCERATION IN THE UNITED STATES: EXPLORING CAUSES AND CONSEQUENCES 46 (Jeremy Travis et al. eds., 2014).

20. I explore the socio-political dynamics shaping U.S. mens rea policy at length in *Strict Liability Abolition*. Serota, *supra* note 6, at 120–41. For earlier interventions on the topic, see Benjamin Levin, *Mens Rea Reform and Its Discontents*, 109 J. CRIM. L. & CRIMINOLOGY 491 (2019); Darryl K. Brown, *Criminal Law Reform and the Persistence of Strict Liability*, 62 DUKE L.J. 285 (2012); Jonathan Simon, *Wechsler's Century and Ours: Reforming Criminal Law in a Time of Shifting Rationalities of Government*, 7 BUFF. CRIM. L. REV. 247 (2003).

21. Letter from Oliver Wendell Holmes, Assoc. Just., Sup. Ct. of the U.S., to Harold J. Laski, Professor, Harv. Univ. (July 14, 1916), *in* 1 HOLMES-LASKI LETTERS: THE CORRESPONDENCE OF MR. JUSTICE HOLMES AND HAROLD J. LASKI 4 (Mark DeWolfe Howe ed., 1953) (emphasis omitted) ("I always have thought that most of the difficulties as to the *mens rea* was due to having no precise understanding what the *mens rea* is."); *see infra* Part I.

community? A growing body of experimental research reveals the multiplicity of mental phenomena that impact people's judgments of moral and criminal responsibility.[22] But this social science literature does not explain how our nuanced intuitions about the minds of others fit together. And it also fails to supply the kind of conceptual roadmap necessary to more closely align criminal policy with psychological blameworthiness.

Perhaps most important, many struggle to understand mens rea's moral significance. For example, in a time of overwhelming prison populations and pervasive racial injustices, why does it matter whether criminal liability is limited to those who act with a guilty mind? Those working the frontlines of the fight against mass incarceration are unlikely to find an accessible explanation in contemporary mens rea scholarship.[23] Nor does this philosophically dense literature offer criminal reformers much in the way of innovative policy solutions: For over sixty years, conversations about mens rea reform have revolved around the Model Penal Code's proposal for universal culpable mental state requirements, with very little thought or consideration given to other legislative pathways.[24]

This Article seeks to breathe new life into mens rea policy and reform by reconstructing a centuries-old understanding of the topic. The "Guilty Minds" approach driving this study is broad and moralistic; it defines mens rea as any and all mental (or quasi-mental) aspects of blame. Undertheorized and poorly administered, this common law conception has largely been replaced instead with the Model Penal Code's reconceptualization of mens rea as purpose, knowledge, recklessness, or negligence. Modern criminal law's embrace of this narrower and more legalistic "PKRN" approach has brought with it many well-known benefits. But there are also overlooked costs of divorcing mens rea doctrine from its moral foundations. This Article demonstrates how the Guilty Minds approach, once clarified and refined, can

22. This experimental research is discussed at length *infra* Part II.

23. The philosophical justifications for mens rea provided by contemporary criminal law theory research, as well as my own gloss on them, are discussed *infra*, Part III. That discussion is part of my broader effort to highlight the salience of mens rea reform in an era of mass incarceration. For example, in another recent piece, *Does Mens Rea Matter?*, I work with a team of social scientists to empirically analyze the legal impact of culpable mental state requirements on charging and conviction rates. *See* Matthew L. Mizel, Michael Serota, Jonathan Cantor & Joshua Russell-Fritch, *Does Mens Rea Matter?*, 2023 WIS. L. REV. (forthcoming 2023) (on file with the *Maryland Law Review*) (internal and external peer review facilitated by the RAND Corporation). And in *Strict Liability Abolition*, I explain why these findings, when viewed in light of the distribution of strict liability in U.S. criminal codes, indicates that universal mens rea requirements would be an effective (if modest) pathway for promoting decarceration and racial justice in a political climate that may otherwise be inhospitable to criminal justice reform. *See* Serota, *supra* note 6.

24. *See* Serota, *supra* note 6, at 120–41 (discussing the history of mens rea reform in the United States).

address these costs and offer all U.S. jurisdictions with a promising new pathway for criminal law reform.

This project unfolds in four parts. Part I unpacks what mens rea is—on the level of both definition and doctrine. After elucidating the concepts most integral to understanding mens rea, this Part outlines the two primary perspectives on mens rea policy: Guilty Minds and PKRN. Although the PKRN approach is widely assumed to be universally superior due to its analytic and administrative strengths, I argue that the Guilty Minds understanding of mens rea as psychological blameworthiness has something important to offer the criminal law. This broad and moralistic conception of mens rea offers a critical lens for understanding why mental states matter, which mental states are morally significant, and where contemporary criminal policies fail to protect actors whose minds do not warrant the condemnation of a criminal conviction.

Part II transforms the vague notion of psychological blameworthiness driving the Guilty Minds approach into an intelligible legal principle. Synthesizing the findings of experimental social science research, this Part develops a model of culpability in which the public's moral assessments of the minds of wrongdoers are centrally influenced by four psychological domains: (1) motivation, (2) risk awareness, (3) rationality, and (4) volitional control. After exploring how each of these mental phenomena contribute to the community's sense of psychological blameworthiness, this Part details how disparate lay intuitions about the minds of others often converge into shared judgments of moral and criminal responsibility.

Part III explains why mens rea, understood in terms of the community's sense of psychological blameworthiness, should constrain the scope of criminal liability in U.S. criminal codes. The key to my argument lies in recognizing what criminal judgments *are*—namely, formalized public expressions of blame—and what this particular class of expressions *says*—namely, that someone failed to care enough about legally-protected individual and societal interests. By understanding the social meaning of blame and the moral significance of mental states, we come to see why psychological blameworthiness provides a critical limit on the scope of liability and punishment in U.S. criminal codes.

Part IV explores the two most important implications of this argument for U.S. criminal policy. First, requiring the government to prove purpose, knowledge, recklessness, or negligence as to every element of an offense is a necessary but inadequate form of mens rea protection. The Guilty Minds culpability model reveals structural flaws in contemporary mens rea policies that authorize convictions for insufficiently blameworthy actors. Second, these flaws should be addressed through a statutory "insufficient blameworthiness" defense, which empowers factfinders to dismiss charges

based on a structured assessment of mitigating mental states, consistent with the Guilty Minds culpability model. Intended to operate in tandem with PKRN mens rea requirements, I explore the legislative details of the proposed defense, discuss important precedents for employing it, and highlight the strong administrative case in support of enacting it.

I. THE LAW OF GUILTY MINDS

Mens rea is both a central preoccupation of American criminal law and a perennial source of confusion. This Part seeks to dispel some of the confusion by integrating multiple perspectives on mens rea into a unified narrative. This narrative includes a philosophical analysis of mens rea terms, a historical analysis of mens rea doctrine, and a functionalist analysis of how different conceptions of mens rea influence the breadth of criminal liability.

Section I.A presents the ideas most integral to understanding the evolution of mens rea in U.S. criminal law. This Section first discusses the meaning of fundamental mens rea concepts, and thereafter provides a historical overview of the two primary perspectives on mens rea reflected in U.S. criminal law: the common law's Guilty Minds approach, and the Model Penal Code's PKRN approach. Typically, these two perspectives are understood to be in conflict with one another; however, this Section argues that they each serve a different purpose—and thus both have something valuable to offer the criminal law. Whereas PKRN's analytic precision supports the production of administrable mens rea doctrine, the moralistic focus of the Guilty Minds approach leads to clearer understanding of why mens rea matters, which mental states are blameworthy, and where contemporary mens rea policies fail to protect actors whose minds do not warrant the condemnation of a criminal conviction.

Section I.B illustrates this distinction, and the practical consequences of conceptualizing mens rea in different ways, through an analysis of a recent District of Columbia case, *Crossland v. United States*.[25] The accused in this case, Mr. Crossland, was convicted of purposely assaulting two police officers who were unlawfully harassing him on racially-discriminatory grounds. Although Mr. Crossland clearly possessed mens rea in the PKRN sense, viewing the case through the lens of the Guilty Minds approach indicates Mr. Crossland's state of mind may be insufficiently blameworthy to merit criminal liability under the circumstances.

25. 32 A.3d 1005 (D.C. 2011).

A. Mens Rea Policy: Key Concepts and Conflicting Perspectives

Mens rea is a central preoccupation of American criminal law.[26] From model codes,[27] to judicial opinions,[28] to scholarship,[29] to casebooks,[30] criminal law discourse in the United States has revolved around the topic for more than a century.[31] What, then, *is* mens rea? It is surprisingly difficult to find a clear answer.

Mens rea incontrovertibly has something to do with the mind. But beyond that, "courts and writers are in hopeless disagreement."[32] "The problem," as one leading casebook author phrases it, "is that no one really knows, or if they know, haven't clearly said," which aspects of the mind are implicated by the term.[33] Absent a precise accounting of mental states, mens rea has developed a distinctive "chameleon-like"[34] quality, taking on different meanings in different legal contexts, in a way that is unrivaled "for

26. *See, e.g.*, Douglas Husak, *"Broad" Culpability and the Retributivist Dream*, 9 OHIO ST. J. CRIM. L. 449, 449 (2012) ("The criminal law is centrally concerned with culpability or mens rea—roughly, the mental (or quasi-mental) components of blame."); Gideon Yaffe, *The Point of Mens Rea: The Case of Willful Ignorance*, 12 CRIM. L. & PHIL. 19, 44 (2018) ("Criminal law exhibits a well-known fixation with the defendant's mind, a fixation that we do not find in other areas of law, including areas in which the mental states of the parties matter to liability."); Gary V. Dubin, *Mens Rea Reconsidered: A Plea for a Due Process Concept of Criminal Responsibility*, 18 STAN. L. REV. 322, 325 (1966) (observing the "nearly deified legal status" of mens rea).

27. MODEL PENAL CODE & COMMENTARIES (AM. L. INST. 1985).

28. *See, e.g.*, Morissette v. United States, 342 U.S. 246, 250–51 (1952); United States v. Cordoba-Hincapie, 825 F. Supp. 485, 492 (E.D.N.Y. 1993).

29. *See, e.g.*, STEPHEN P. GARVEY, GUILTY ACTS, GUILTY MINDS (2020); GIDEON YAFFE, THE AGE OF CULPABILITY: CHILDREN AND THE NATURE OF CRIMINAL RESPONSIBILITY (2018); DOUGLAS HUSAK, THE PHILOSOPHY OF CRIMINAL LAW: SELECTED ESSAYS (2010); LARRY ALEXANDER & KIMBERLY KESSLER FERZAN, CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW (William A. Edmundson & Brian Bix eds., 2009); H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW (2d ed. 2008); Peter Westen, *An Attitudinal Theory of Excuse*, 25 LAW & PHIL. 289 (2006); GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW (2000); MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW (1997); Sanford H. Kadish, *Excusing Crime*, 75 CALIF. L. REV. 257 (1987); JEROME HALL, GENERAL PRINCIPLES OF CRIMINAL LAW (2d ed. 1960); Sayre, *supra* note 1.

30. *See, e.g.*, JOSHUA DRESSLER & STEPHEN P. GARVEY, CRIMINAL LAW: CASES AND MATERIALS (8th ed. 2019); SANFORD H. KADISH, STEPHEN J. SCHULHOFER & RACHEL E. BARKOW, CRIMINAL LAW AND ITS PROCESSES: CASES AND MATERIALS (10th ed. 2016); PAUL H. ROBINSON, SHIMA BARADARAN BAUGHMAN & MICHAEL T. CAHILL, CRIMINAL LAW: CASE STUDIES AND CONTROVERSIES (5th ed. 2020).

31. *See, e.g.*, Claire Finkelstein, *The Inefficiency of Mens Rea*, 88 CALIF. L. REV. 895, 896 (2000) ("[T]he criminal law's requirement of mens rea is the central distinguishing characteristic of the institution.").

32. Sayre, *supra* note 1, at 974.

33. Stephen P. Garvey, *Authority, Ignorance, and the Guilty Mind*, 67 SMU L. REV. 545, 546 (2014).

34. Francis Bowes Sayre, *The Present Signification of* Mens Rea *in the Criminal Law*, *in* HARVARD LEGAL ESSAYS 399, 402 (1934).

the quantity of obfuscation it has created."[35] Remarking on these definitional issues, one prominent criminal law scholar concludes: "There is no term fraught with greater ambiguity . . . ."[36]

Ambiguous though the term may be, let us try to make some definitional headway. Roughly translated, mens rea means "guilty mind" in Latin.[37] The significance of this term is derived from the oft-cited Latin phrase "[*a*]*ctus non facit reum nisi mens rea*," which means "[t]he act is not culpable unless the mind is guilty."[38] The U.S. legal system has its own version of this axiom, which the U.S. Supreme Court first articulated in *Morissette v. United States*[39] and U.S. legal culture has fervently recounted ever since: "Crime, as a compound concept, generally [requires the] concurrence of an evil-meaning mind with an evil-doing hand . . . ."[40] The basic idea behind these aphorisms is intuitive: It is always regrettable when people do the wrong thing, but the criminal legal system should only blame people—through its formalized methods of conviction and punishment—when their wrongful conduct is accompanied by a culpable state of mind.

This principle of mens rea is built upon a few philosophically slippery concepts that arise at the intersection of moral psychology and criminal law theory.[41] The most central is blame,[42] which is a judgment of disapproval of

35. Sanford H. Kadish, *The Decline of Innocence*, 26 CAMBRIDGE L.J. 273, 273 (1968).

36. FLETCHER, *supra* note 29, at 398.

37. *Mens Rea*, BLACK'S LAW DICTIONARY 1181 (11th ed. 2019).

38. Garvey, *supra* note 33, at 545, 545 n.1 (observing that "[t]he phrase is usually traced to St. Augustine's Sermon 180 on perjury" (citing Saint Augustine, Sermon 180, *in* 5 THE WORKS OF SAINT AUGUSTINE: A TRANSLATION FOR THE 21ST CENTURY, PT. III—SERMONS 314, 315 (John E. Rotelle ed., Edmund Hill trans., New York City Press 1992))); *see also* 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 21 (noting that "an unwarrantable act without a vitious will is no crime at all").

39. 342 U.S. 246 (1952).

40. *Id.* at 251.

41. David O. Brink, *The Nature and Significance of Culpability*, 13 CRIM. L. & PHIL. 347, 347–48 (2019) (observing "the importance of moral ideas to the formation and reform of criminal law principles and practices and the effect of well-settled criminal law doctrine on our moral assumptions and beliefs," and noting the "mutual influence and interaction between these domains").

42. This Article only endeavors to provide a simple, but useful, working understanding of blame that serves the purpose of this Part. In reality, the concept of blame is just as complex as that of mens rea. David Shoemaker, *Blameworthy but Unblamable: A Paradox of Corporate Responsibility*, 17 GEO. J.L. & PUB. POL'Y 897, 905 (2019). This complexity is reflected in the robust philosophical literature on the nature of blame, which has produced a number of understandings, including that:

- Blame is a reactive attitude, such as anger, resentment, and indignation. Peter F. Strawson, *Freedom and Resentment*, *in* FREE WILL 72, 72–73 (Gary Watson ed., 2d ed. 2003); *see* R. JAY WALLACE, RESPONSIBILITY AND THE MORAL SENTIMENTS 75 (1994) (stating that it would be "strange to suppose that one might blame another person without feeling an attitude of indignation or resentment toward the person").

a human action or quality,[43] as well as a pervasive social practice.[44] Most often, we blame people for the (bad) things they do, which brings with it a judgment of responsibility for the negative effects caused or threatened by their conduct.[45]

The social practice of blaming generally rests upon two main evaluative criteria: wrongfulness and culpability. Wrongfulness is typically understood to be an attribute of *acts*,[46] whereas culpability is an attribute of *actors*.[47] A person's act is wrongful when (and to the extent) it violates an accepted norm

- Blame is a pair of mental activities, namely, (1) a judgment that someone has done something reflecting "attitudes toward others that impairs the relations that others can have with him or her"; and (2) a modification of one's own attitudes, intentions, or dispositions in a way appropriate to that relationship impairment. T.M. SCANLON, MORAL DIMENSIONS: PERMISSIBILITY, MEANING, BLAME 128–29 (2008).
- Blame is an attitudinal or behavioral response that repudiates or takes a stand against someone's immoral treatment, and is often accompanied by protest that, *inter alia*, aims to get the offender and others in the moral community to acknowledge what the offender did. Angela M. Smith, *Moral Blame and Moral Protest*, *in* BLAME: ITS NATURE AND NORMS 27, 41–47 (D. Justin Coates & Neal A. Tognazzini eds., 2013).
- Blame "signal[s] the blamer's commitment to a set of norms." David Shoemaker & Manuel Vargas, *Moral Torch Fishing: A Signaling Theory of Blame*, 55 NOÛS 581, 581 (2021).

For a recent edited volume that provides a good overview of the field, see BLAME: ITS NATURE AND NORMS, *supra*.

43. Sanford H. Kadish, *Complicity, Cause and Blame: A Study in the Interpretation of Doctrine*, 73 CALIF. L. REV. 323, 329 (1985) ("Blame and its correlative, praise, serve as expressions of our disapproval or approval of some human action or quality."); *id.* ("[Blame] is not intrinsically an expressive action, but a judgment of disapproval. It is an internal evaluation that need not be expressed. Blame is the sentiment of disapproval itself.").

44. *Id.* ("Attributing blame is a pervasive human phenomenon. It is one way in which we order and make sense of social experience and it is reflected in our language and social practices.").

45. *Id.* at 330 ("[B]lame entails a judgment of responsibility . . . . One who presents a poor appearance is not blamed for it unless he is responsible for it, as by poor judgment or lack of care in dress or grooming. The runner who loses a race may be blamed if he failed to train properly, but not if he did the best he could."); *see, e.g.*, Meir Dan-Cohen, *Responsibility and the Boundaries of the Self*, 105 HARV. L. REV. 959, 959 (1992) ("[Blame involves] ascribing moral responsibility for the negative effects of one's behavior . . . .").

46. *See, e.g.*, YAFFE, *supra* note 29, at 66 ("Wrongfulness is a property of actions."); Peter Westen, *Individualizing the Reasonable Person in Criminal Law*, 2 CRIM. L. & PHIL. 137, 151 (2008); Kadish, *supra* note 43, at 329–32.

47. *See, e.g.*, YAFFE, supra note 29, at 68 ("Culpability is not, like wrongfulness, a property of acts, but, instead, a property of agents. However, agents are not culpable, full stop. Rather, they are culpable for acts, and not act types, but act tokens."); Peter Westen, *Unwitting Justification*, 55 SAN DIEGO L. REV. 419, 420 (2018) ("It is commonplace, for example, to conceptualize criminal responsibility as consisting of two elements—wrongdoing and culpability—where culpability, like *mens rea*, consists exclusively of an actor's mental capacities and states, and wrongdoing, like *actus reus*, consists of something distinct from that.").

of conduct.[48] Many categories of these behavioral norms exist, ranging from the aesthetic (e.g., "color within the lines") to the professional (e.g., "write criminal law papers clearly").[49] Most integral to the practice of blaming, however, is the category of moral norms, which implicate the well-being and interests of other people (e.g., "don't force people to read unclear criminal law papers").[50]

Conduct that violates a moral norm is morally wrongful.[51] And a finding of morally wrongful conduct is critical to most expressions of blame.[52] While critical, however, morally wrongful conduct is not independently sufficient. Our blaming practices are also predicated on another evaluative criterion, culpability, which focuses on the mind of the person (in contrast to the bad thing they did).[53] Specifically, culpability is comprised of those mental characteristics that enable morally wrongful conduct to be attributed to the actor in a way that makes blame fitting or appropriate (or which mediates between degrees of blameworthiness).[54]

What remains to be determined, not only for moral philosophy but also for criminal policy, is: (1) which mental characteristics constitute culpability (i.e., contribute to or detract from blame); and (2) how should normative assessments of the minds of wrongdoers be structured for those who sit in judgement of others? These basic questions animate a longstanding debate over the meaning and structure of mens rea policy. As discussed below, two main perspectives have been brought to bear on the topic: a broader and more

48. *See, e.g.*, YAFFE, *supra* note 29, at 66 ("An action is wrongful (or, equivalently, an instance of wrongdoing) just in case it is in violation of a norm; it is not-to-be-done."); Westen, *supra* note 46, at 151; Kadish, *supra* note 41, at 329–32.

49. *See, e.g.*, YAFFE, *supra* note 29, at 67 ("[There are different] sets of norms. If you're painting a portrait, it might be wrong to paint the eyes blue. It is not morally wrong, but aesthetically wrong, or maybe wrong because it violates norms of a particular genre of painting, such as realistic portraiture." (emphasis omitted)); Westen, *supra* note 46, at 151; Kadish, *supra* note 43, at 329–32.

50. *See, e.g.*, YAFFE, *supra* note 29, at 67 ("Typically, the relevant kind of wrongfulness is moral wrongfulness and so the relevant set of norms are moral norms."); Westen, *supra* note 46, at 151; Kadish, *supra* note 43, at 329–32; *see also* Heidi M. Hurd, *Justification and Excuse, Wrongdoing and Culpability*, 74 NOTRE DAME L. REV. 1551, 1558 (1999) ("Moral wrongdoing consists of doing an action that violates the maxims of our best moral theory—whatever that theory may be, be it consequentialist or deontological.").

51. *See, e.g.*, YAFFE, *supra* note 29, at 66–68; Kadish, *supra* note 43, at 329–32.

52. *See, e.g.*, YAFFE, *supra* note 29, at 68 ("To say that someone is 'culpable for an act' is to imply two things: he is responsible for the act, and the act is wrongful. So culpability is a species of responsibility: it is responsibility for wrongful action."); Westen, *supra* note 46, at 151; Kadish, *supra* note 43, at 329–32.

53. *See, e.g.*, YAFFE, *supra* note 29, at 68; Kadish, *supra* note 43, at 329–32.

54. *See, e.g.*, Westen, *supra* note 46, at 151; Kadish, *supra* note 43, at 329–32; Husak, *supra* note 26, at 454–59.

moralistic Guilty Minds approach, and a narrower and more legalistic PKRN approach.[55]

1. The Guilty Minds Approach

The Guilty Minds approach understands mens rea broadly, to encompass all mental characteristics that contribute to blameworthiness (what this Article refers to a "psychological blameworthiness").[56] This expansive view of mens rea is rooted in the common law and has historically been implemented through holistic, and largely unstructured, psychological assessments of people who violate criminal prohibitions.[57] These assessments seek to generally determine whether a wrongdoer is "blameworthy in mind,"[58] without clearly delineating which particular mental characteristics so qualify. On this broad accounting of mens rea, any psychological phenomena indicating the presence (or absence) of "bad character, malevolence, or immorality"[59] could be considered to fall within the purview of legal decision-making.

Elizabeth Papp Kamali's recent study of mens rea policy in fourteenth-century England provides a helpful illustration. During this era, medieval courts and juries were apt to consider a criminal defendant's motives, emotions, and sanity along with the influence of external coercion and cognitive deficiencies in adjudicating criminality.[60] That evaluation, as

55. *See, e.g.*, Husak, *supra* note 26, at 449–50; Kadish, *supra* note 35, at 274–75; JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW § 10.02 (8th ed. 2018); Serota, *supra* note 7, at 1202–03 n.4. As Steve Garvey explains, this distinction has been framed in a number of ways—for example, he characterizes it as the choice between "mens rea and *mentes reae*," while others describe it as the "distinction between culpability and elemental accounts of mens rea, or normative and descriptive accounts, or broad and narrow accounts." Garvey, *supra* note 33, at 546–47 n.2. I am reticent to add yet another framing of this distinction; however, the different terminology employed in this article—"PKRN" and "Guilty Minds"—is more descriptively accurate, and thus may be more likely to ease the cognitive burden on readers who are not already familiar with this debate.

56. Husak, *supra* note 26, at 449; DRESSLER, *supra* note 55, § 10.02[B].

57. For insightful discussions of this common law tradition, see KAMALI, *supra* note 1; DRESSLER, *supra* note 55, § 10.02; Martin R. Gardner, *The* Mens Rea *Enigma: Observations on the Role of Motive in the Criminal Law Past and Present*, 1993 UTAH L. REV. 635.

58. Morissette v. United States, 342 U.S. 246, 252 (1952); *see* Kahler v. Kansas, 140 S. Ct. 1021, 1042 (2020) ("At common law, the term *mens rea* ordinarily incorporated the notion of 'general moral blameworthiness' required for criminal punishment. . . . as used at common law, the term *mens rea* 'is synonymous with a person's blameworthiness.'[]" (first quoting Sayre, *supra* note 1, at 988; and then quoting JOSHUA DRESSLER, 3 ENCYCLOPEDIA OF CRIME & JUSTICE 995 (2d ed. 2002))); STUNTZ, *supra* note 18, at 260 ("Traditionally, [the law of mens rea] required proof that the defendant acted with a state of mind that was worthy of moral blame.").

59. DRESSLER, *supra* note 55, § 10.02[B].

60. KAMALI, *supra* note 1, at 86–88.

Kamali describes it, appears to have been both *simultaneous* and *graduated*.[61] That is, in assessing mens rea, decisionmakers would consider narrow mental characteristics such as intentionality—for example, whether a given harm was caused premeditatedly, deliberately but without planning, or carelessly—alongside broader psychological phenomena, such as a criminal defendant's reasons for acting and their ability to make rational decisions.[62] Within these evaluations of psychological blameworthiness, decisionmakers seemed to subject disparate mental states to a sliding scale analysis that accounted for the mitigating or aggravating strength of each in relation to one another.[63]

These holistic mens rea assessments, while nuanced and contextual, were largely unregulated by the strictures of crystallized doctrine. That is because criminal statutes characteristic of the common law era tend to offer little more than vague references to a criminal offender's "felonious intent," "ill will," or innumerable other undefined mental state terms.[64] Absent a clear legal definition or explicit legal guidance, decisionmakers would rely on their own sense justice in conducting evaluations of psychological blameworthiness.

This same basic dynamic holds in jurisdictions that continue to employ vague common law mens rea terminology. From murder prosecutions premised on an offender's "depraved heart,"[65] to manslaughter prosecutions that focus on an offender's "wanton[ness],"[66] to property destruction prosecutions centered around an offender's "malice,"[67] many statutes and judicial opinions continue to conceptualize mens rea in broad normative

61. *See id.* at 87 ("Most assuredly medieval juries did not have a checklist to assess a defendant's intentionality, freedom to act, and rationality, yet these factors appear often enough in trial evidence to suggest that they were among the multiplicity of norms that influenced jury decision-making."); *id.* at 86 ("In assigning punishment, the intent of the accused determined the severity. Intentionality mattered in weighing crime and sin even prior to the advent of jury trial for felony.").

62. KAMALI, *supra* note 1, at 87–88. For example, as Martin Gardner explains, on this accounting "not only must an offender's acts be intended but his ulterior motives or purposes in acting must also be blameworthy." Gardner, *supra* note 57, at 658. And mens rea is similarly understood to "require[] that offenders function as moral agents rationally choosing their evil designs." *Id.* at 662. Viewed through this lens, mens rea "constitute[s] a normative judgment of subjective wickedness, requiring not simply that the actor intend to commit the offense, but also that the offense be committed by a responsible moral agent for wicked purposes." *Id.* at 663.

63. *See* KAMALI, *supra* note 1, at 87–88.

64. Gardner, *supra* note 57, at 669 n.178, 672 n.193, 676 n.211 (quoting PETER W. LOW, JOHN CALVIN JEFFRIES, JR. & RICHARD J. BONNIE, CRIMINAL LAW: CASES AND MATERIALS 198–99 (2d ed. 1986)); *see* Paul H. Robinson & Jane A. Grall, *Element Analysis in Defining Criminal Liability: The Model Penal Code and Beyond*, 35 STAN. L. REV. 681, 692 (1983) (observing the existence of approximately eighty mental state terms at common law).

65. *E.g.*, Windham v. State, 602 So. 2d 798, 800 (Miss. 1992).

66. *E.g.*, State v. Albrecht, 336 Md. 475, 486–87, 649 A.2d 336, 341 (1994).

67. *E.g.*, Brown v. United States, 584 A.2d 537, 539 (D.C. 1990).

terms. However, this common law conceptualization is also being applied in increasingly fewer situations, and with increasingly greater limitations on the exercise of juror discretion.[68]

The gradual retrenchment of the Guilty Minds approach is not surprising. Delegating large swaths of moral discretion to decisionmakers risks inconsistent and arbitrary outcomes that are anathema to contemporary criminal law, which is founded on the legality principle[69] and prioritizes second order values such as fair notice, equality, and predictability.[70] These values are frustrated by inviting juries to undertake broad, unstructured assessments of psychological blameworthiness, which create a breeding ground for cognitive bias, arbitrariness, and inconsistent outcomes.[71]

Similar problems arise from asking courts to fill in the meaning of vague culpability terms in the context of adjudicating individual cases.[72] During the first half of the twentieth century, this common law method of policy creation infamously produced an "amorphous . . . quagmire" of mens rea doctrine comprised of "a thin surface of general terminology denoting wrongfulness."[73] It thus became increasingly clear during the second half of the twentieth century that a change was in order.

2. The PKRN Approach

The drafters of the Model Penal Code sought to bring about that change through a more analytically precise and legalistic understanding of mens rea.[74] What they produced, the PKRN approach, conceives of mental state

68. *See also* STUNTZ, *supra* note 18, at 260 (observing that while "[s]ome vestiges" of the common law approach to mens rea exist, "for the most part, the concept of wrongful intent—the idea that the state must prove the defendant acted with a 'guilty mind,' . . . —has gone by the boards").

69. *See, e.g.*, John Calvin Jeffries, Jr., *Legality, Vagueness, and the Construction of Penal Statutes*, 71 VA. L. REV. 189 (1985); Alan C. Michaels, *"Rationales" of Criminal Law Then and Now: For a Judgmental Descriptivism*, 100 COLUM. L. REV. 54, 72–73 (2000).

70. *See* Serota, *supra* note 7, at 1210–11 (discussing the influence of the legality principle on criminal law, and the values it embodies).

71. *See, e.g.*, Michael Serota, *Mens Rea, Criminal Responsibility, and the Death of Freddie Gray*, 114 MICH. L. REV. FIRST IMPRESSIONS 31 (2015); Paul H. Robinson, *Legality and Discretion in the Distribution of Criminal Sanctions*, 25 HARV. J. ON LEGIS. 393 (1988).

72. *See* Serota, *supra* note 7, at 1214 (absent clear legislative statements about mental state requirements, common law "judges often struggled mightily to resolve the relevant mens rea issues, disparately applying judicially created policies that were themselves 'inconsistent and confusing.'" (quoting LOW ET AL., *supra* note 64, at 198–99)).

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

74. *See, e.g.*, Herbert Wechsler, *Codification of Criminal Law in the United States: The Model Penal Code*, 68 COLUM. L. REV. 1425 (1968). For an early recognition of this framing, see Sayre, *supra* note 34, at 404 ("An intelligent understanding of the various states of mind requisite for criminality can be gained only through an intensive study of the substantive law covering each

evaluations as a predominantly descriptive enterprise made up of discreet inquiries, clearly specified by statute, that target particular aspects of a criminal defendant's state of mind.[75]

At the heart of this enterprise is an assessment of whether a criminal defendant possessed one of four culpable mental states—what most jurisdictions refer to as purpose, knowledge, recklessness, and negligence (hence, "PKRN"). Here is a simplified overview of how these common legislative terms are understood to apply in contemporary U.S. criminal law:

(1) A person acts *purposely* when it is their conscious object to engage in wrongdoing[76]—i.e., to cause a prohibited result or to commit a prohibited act under specified circumstances.[77] Purpose is reflected where, for example, *D* pulls the trigger of a loaded gun with the goal of killing *V*. If *D* is successful, then *D* purposely killed *V*.

(2) A person acts *knowingly* when they possess a high level of awareness of wrongdoing[78]—i.e., that their conduct would cause a prohibited result or that it occurred under prohibited circumstances.[79] Knowledge is reflected where, for example, *D*, a child rights advocate, blows up a manufacturing facility that relies on youth labor, aware that *V*, the on-duty night guard who is

separate group. The old conception of *mens rea* must be discarded, and in its place must be substituted the new conception of *mentes reae*.").

75. *See, e.g.*, Michaels, *supra* note 69, at 58 (observing that the Model Penal Code drafters embrace a "descriptive approach to the criminal law" understood in terms of "a preference for seeking to identify the facts that ought to be determinative of liability rather than relying on vaguer standards that call more for subjective appraisal and assessment than for precise fact finding"); Vera Bergelson, *The Depths of Malice*, 53 ARIZ. ST. L.J. 399, 399 (2021) (observing the Model Penal Code's "strive for simplification, rationality, and utility" in the law of mens rea); Gardner, *supra* note 57, at 688 (observing that the Model Penal Code "treats mental offense elements essentially as descriptive states of mind"). For discussion of two important exceptions to this characterization of the Model Penal Code's codification philosophy, see *infra* IV.B.1–2 (discussing the Model Penal Code's gross deviation standard—for recklessness and negligence assessments—and its *de minimis* defense).

76. *See* MODEL PENAL CODE § 2.02(2)(a) (AM. L. INST., Proposed Official Draft 1962).

77. That wrongdoing is one's conscious object is to be distinguished from wrongdoing that is *voluntarily* committed. *See, e.g.*, Paul H. Robinson, *A Functional Analysis of Criminal Law*, 88 NW. U. L. REV. 857, 864 (1994) (explaining that the voluntariness requirement, which implicates what is sometimes referred to as a "*present* conduct intention," can be "satisfied simply by showing that the actor did in fact intend to perform the bodily movements that he performed"). The mere fact that someone engages in wrongdoing voluntarily is entirely consistent with strict liability. *See, e.g.*, Buchanan v. United States, 32 A.3d 990, 1002 (D.C. 2011) (Ruiz, J., concurring) (noting that the "intent to commit" interpretation of simple assault, if taken literally, "would allow the prosecution of individuals for criminal assault for actions taken with a complete lack of culpability").

78. *See* MODEL PENAL CODE § 2.02(2)(b) (AM. L. INST., Proposed Official Draft 1962).

79. The reference to "awareness of wrongdoing," both here and throughout the Article, does not implicate awareness that one's conduct might be illegal—although that too could implicate the blameworthiness of a person's state of mind. *See, e.g.*, DOUGLAS HUSAK, IGNORANCE OF LAW: A PHILOSOPHICAL INQUIRY (2016).

unaffiliated with the business, is practically certain to die in the blast. If *V* does die in the blast, then the law says that *D* knowingly killed *V*.

(3) A person acts *recklessly* when (among other requirements) they consciously disregard a substantial risk of wrongdoing[80]—i.e., that their conduct would cause a prohibited result or that it occurred under prohibited circumstances. Recklessness is reflected where, for example, *D* speeds through a red light aware that it is substantially possible that their car will fatally hit *V*, a pedestrian stepping into the crosswalk. If *V* is fatally struck, then *D* is likely to have recklessly killed *V*.[81]

(4) A person acts *negligently* when (among other requirements) they fail to be aware of a substantial risk of wrongdoing[82]—i.e., that their conduct would cause a prohibited result or that it occurred under prohibited circumstances—where a reasonable person, if placed in the same situation, would have perceived that risk. Negligence is reflected where, for example, *D* speeds through a red light while reading their phone, unaware that it is substantially possible that their car will fatally hit *V*, a pedestrian stepping into the crosswalk. If *V* is fatally struck, then *D* is likely to have negligently killed *V*.[83]

In developing the PKRN approach, the drafters of the Model Penal Code believed that these four clearly defined culpable mental states were all that legislatures "needed to prescribe the minimal requirements [for liability] and lay the basis for [grading] distinctions that may usefully be drawn."[84] Judging from widespread adoption of the Model Penal Code's reconceptualization of mens rea, most legal authorities seem to agree.[85] For example, a strong majority of the thirty-five states that successfully modernized their codes

80. *See* MODEL PENAL CODE § 2.02(2)(c) (AM. L. INST., Proposed Official Draft 1962).

81. I say "likely" because the risk consciously disregarded by the accused must also be unjustified and the accused's conduct must also satisfy the gross deviation standard governing recklessness inquiries. *See* MODEL PENAL CODE § 2.02(2)(c) (AM. L. INST., Proposed Official Draft 1962). For further discussion of the gross deviation standard, see *infra* IV.B.1.

82. *See* MODEL PENAL CODE § 2.02(2)(d) (AM. L. INST., Proposed Official Draft 1962).

83. Here again I say "likely" because the risk that the accused failed to perceive must have been unjustified and the accused's conduct must also satisfy the gross deviation standard governing negligence inquiries. *See* MODEL PENAL CODE § 2.02(2)(d) (AM. L. INST., Proposed Official Draft 1962). For further discussion of the gross deviation standard, see *infra* IV.B.1.

84. Wechsler, *supra* note 74, at 1436.

85. *See, e.g.*, Paul H. Robinson & Markus D. Dubber, *The American Model Penal Code: A Brief Overview*, 10 NEW CRIM. L. REV. 319 (2007) (describing legislative trends); D.C. CRIM. CODE REFORM COMM'N, RECOMMENDATIONS FOR CHAPTER 2 OF THE REVISED CRIMINAL CODE: BASIC REQUIREMENTS OF OFFENSE LIABILITY (2016), https://ccrc.dc.gov/sites/default/files/dc/sites/ccrc/publication/attachments/Report-2-First-Draft-Recommendations-for-Chapter-2-of-the-Revised-Criminal-Code-Basic-Requirements-of-Offense-Liability.pdf (summarizing the influence of the PKRN approach on U.S. criminal law).

during the latter half of the twentieth century adopted some version of the Model Penal Code's culpable mental state hierarchy, while jurisdictions that failed to overhaul their codes still frequently rely on that hierarchy in creating mens rea policy.[86] As a result, the PKRN approach is now understood to be the "representative modern American culpability scheme,"[87] as well as "a standard part of the furniture of the criminal law."[88] Consequently, insofar as a criminal defendant's state of mind is at issue in prosecutions arising under modern criminal codes, the search for PKRN comprises nearly all of the focus.

I say "nearly" because mental state evaluations in both common law and Model Penal Code jurisdictions are supplemented by affirmative defenses that invite focused consideration of other aspects of an actor's state of mind. Consider, for example, three of the most prevalent:

(1) A person who, by virtue of serious mental illness, is unable to appreciate the criminality of their conduct or conform their conduct to the requirements of law may raise an insanity defense in order to avoid criminal liability.[89]

(2) A person who, by virtue of an unlawful threat of imminent death or severe bodily injury to themself or another, is coerced into

86. Serota *supra* note 6, at 129–30, 129–30 nn. 95–96 (collecting relevant authorities).

87. Robinson & Grall, *supra* note 64, at 692.

88. Francis X. Shen et al., *Sorting Guilty Minds*, 86 N.Y.U. L. REV. 1306, 1318 (2011) (quoting Sanford H. Kadish, *The Model Penal Code's Historical Antecedents*, 19 RUTGERS L.J. 521, 521 (1988)).

89. Generally speaking, the excuse of insanity asserts the absence of blameworthiness due to defects of reasoning and moral judgment that were the result of a mental disease which precluded the person from complying with the law. Kadish, *supra* note 29, at 262–63. There are a number of approaches to insanity; however, the *M'Naghten* and Model Penal Code approaches are most common today. Paul H. Robinson et al., *The American Criminal Code: General Defenses*, 7 J. LEGAL ANALYSIS 37, 77–78 (2015) (providing a detailed overview of prevailing legal trends); *see, e.g.*, WAYNE R. LAFAVE, SUBSTANTIVE CRIMINAL LAW, pt. 2, § 7.1(a) (3d ed. 2018 & Supp. 2021); MODEL PENAL CODE § 4.01(1) (AM. L. INST., Proposed Official Draft 1962). The *M'Naghten* test asks whether, as a result of mental disease, the defendant was unable to know the nature of his act, or that the act was wrong. *See, e.g.*, LAFAVE, *supra* ("[U]nder the prevailing *M'Naghten* rule (sometimes referred to as the right-wrong test) the defendant cannot be convicted if, at the time he committed the act, he was laboring under such a defect of reason, from a disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, as not to know he was doing what was wrong."). The Model Penal Code approach expands *M'Naghten* to incorporate a volitional prong—i.e., whether a mental disease prevented the defendant from conforming his conduct to the requirements of the law. MODEL PENAL CODE § 4.01(1) (AM. L. INST., Proposed Official Draft 1962) ("A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of law." (alteration in original)).

committing a criminal offense may raise a duress defense in order to avoid criminal liability.[90]

(3) A person who, motivated by a reasonable belief that they are in immediate danger of unlawful bodily harm from an adversary and that the use of such force is necessary to avoid this danger, commits a crime for defensive purposes may raise self-defense in order to avoid criminal liability.[91]

This Article will have more to say about these affirmative defenses later on.[92] The important thing to note about them for now is simply this: Affirmative defenses at both common law and under the Model Penal Code tend to be narrowly circumscribed (in terms of which mental states they invite consideration of), exceptionally demanding (in terms of how strong the influence of the qualifying mental states needs to be), and therefore limited in the range of circumstances to which they apply.[93] Affirmative defenses are,

90. Generally speaking, the excuse of duress asserts the absence of blameworthiness because a crime was committed under the command of another backed by threats of injury under circumstances in which a person of reasonable fortitude would have done the same. Kadish, *supra* note 29, at 261. There are two main approaches to duress: the common law approach, Robinson et al., *supra* note 89, at 88, and the Model Penal Code approach, MODEL PENAL CODE § 2.09(1) (AM. L. INST., Proposed Official Draft 1962). The common law approach applies in a prosecution for an offense other than murder, where "the actor engaged in the proscribed conduct because he was coerced to do so by what he reasonably believed was an unlawful threat of imminent death or severe bodily injury to himself or another." Robinson et al., *supra* note 89, at 88. The Model Penal Code approach is broader, applying to any offense that was committed because the defendant "was coerced to do so by the use of, or a threat to use, unlawful force against his person or the person of another, which a person of reasonable firmness in his situation would have been unable to resist." MODEL PENAL CODE § 2.09(1) (AM. L. INST., Proposed Official Draft 1962).

91. Generally speaking, the justification of self-defense asserts that a "nonaggressor is justified in using force upon another if he reasonably believes such force is necessary to protect himself from imminent use of unlawful force by the other person." DRESSLER, *supra* note 55, § 18.01; *see, e.g.*, Robinson et al., *supra* note 89, at 49; MODEL PENAL CODE § 3.04(1) (AM. L. INST., Proposed Official Draft 1962). Although self-defense is a justification, where the reasonable beliefs required by the defense are mistaken, it is best understood as an excuse, in that exculpation would be provided on the basis of the absence of blameworthiness. Robinson et al., *supra* note 89, at 75; *see* PAUL H. ROBINSON, 2 CRIMINAL LAW DEFENSES § 184 (2021) (mistake as to a justification); Paul H. Robinson & John M. Darley, *Testing Competing Theories of Justification*, 76 N.C. L. REV. 1095, 1101 (1998) (under these circumstances the "person's conduct is objectively unjustified but the person subjectively, mistakenly believes that it is justified").

92. *See infra* Section IV.A.

93. As highlighted *supra* notes 89–91, the Model Penal Code approach to affirmative defenses such as insanity and duress expanded upon the common law approach to dealing with similar issues. However, even where U.S. legislatures have adopted these expansions, this characterization of affirmative defenses—as narrowly circumscribed, exceptionally demanding, and limited in their applicability—remains accurate. This point is developed further *infra* IV.A.

therefore, complementary but ultimately extrinsic to both the Guilty Minds and PKRN conceptions of mens rea.[94]

3. Mens Rea Concepts and their Functions

Given U.S. criminal law's widespread embrace of the PKRN approach, one can ask: Is the Model Penal Code's reconceptualization of mens rea incontrovertibly better than the common law approach it has come to replace? If the last six decades of criminal scholarship is any guide, the answer seems to be a resounding yes.[95] More recently, however, alternative academic perspectives have surfaced indicating that the situation may be more complicated than it seems.[96] And I tend to agree. What constitutes a "better" approach to mens rea is dependent, at least in part, on the basis for evaluation.

From an administrative perspective, the PKRN approach undoubtedly offers clear advantages over the Guilty Minds approach, which we can see reflected in legal practice. Legislative adoption of the Model Penal Code's culpable mental state framework in reform jurisdictions has yielded a simpler and more consistent body of criminal law, reduced litigation over vague terminology, and afforded defendants better notice.[97] It has also made it substantially easier for practicing lawyers and courts to determine what the prosecution must prove beyond a reasonable doubt.[98] Looking beyond criminal practice, the PKRN approach has also incontrovertibly improved the

94. Section IV.B, *infra*, proposes an affirmative defense centrally concerned with psychological blameworthiness that would change this characterization. Here, however, my analysis focuses on defense law, as presently constituted.

95. *See, e.g.*, Shen et al., *supra* note 88, at 1316 (arguing that the PKRN framework accomplished "what no legal system had ever expressly tried to do: orchestrate the noise of culpability into a reasonably uniform and workable system"); Simon, *supra* note 20, at 248 (observing that the Model Penal Code "since the 1960s has served as the most influential source of modern criminal law reform thought for American scholars and state legislatures").

96. For a notable exception to the general consensus, see STUNTZ, *supra* note 18, at 260–62 (comparing the moral blame-focused mens rea analysis in *Morissette* with the general intent-to-do-harm analysis in *People v. Stark*, 31 Cal. Rptr. 2d 887 (Dist. Ct. App. 1994), and arguing that the PKRN approach "make[s] guilty pleas easier to extract"); *see also* Bergelson, *supra* note 75, at 399 (arguing that, "in its strive for simplification, rationality, and utility, the [Model Penal Code] has sacrificed some of the moral complexity of the traditional, common-law mens rea categories," and using the common law mental state of malice as an illustration); Michaels, *supra* note 69 (comparing descriptive, non-descriptive, utilitarian, and moral blame-focused mens rea formulations in the common law and Model Penal Code approach to homicide, and arguing for judgmental descriptivism); Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV. 463, 488–89 n.89 (1992) (comparing the advantages of "local" and "global" definitions of mental states).

97. Robinson & Grall, *supra* note 64, at 704–06.

98. Dannye Holley, *The Influence of the Model Penal Code's Culpability Provisions on State Legislatures: A Study of Lost Opportunities, Including Abolishing the Mistake of Fact Doctrine*, 27 SW. U. L. REV. 229, 230–31 (1997).

quality of academic and jurisprudential discourse over mens rea by making the conversation more precise, specific, and uniform.[99]

Therefore, insofar as the concept of mens rea exists to segregate, articulate, and organize policy choices about the minds of wrongdoers, there is no question that the PKRN approach is a vast improvement over the Guilty Minds approach. But clarity, precision, and predictability are not the only criteria by which legal concepts are to be evaluated. A legal concept's impact on the content and consequences of government decisions matter, too. And in this regard, conceptualizing mens rea solely in terms of purpose, knowledge, recklessness, or negligence can create problems of its own.

One of the few scholars to recognize this possibility is William Stuntz. In his book, *The Collapse of American Criminal Justice*, Stuntz argues that widespread embrace of the PKRN approach has eroded key protections for criminal defendants and afforded prosecutors greater leverage in plea negotiations.[100] The problem, in Stuntz's view, is that the PKRN approach swaps the common law's broad concern for psychological blameworthiness with a legalistic search for particularized mental states shorn of moral content.[101] This transition, from normativity to technicality, severely constrains the ability of people charged with crimes to raise arguments about the morality of their behavior.[102] As a result, Stuntz concludes, prosecutors operating in a world of PKRN will confront fewer issues that might lead to jury trials and find it easier to extract guilty pleas.[103]

Stuntz's uncompromising assessment of the Model Penal Code approach to mens rea is controversial.[104] Whether, in its totality, the shift from Guilty Minds to PKRN has resulted in a net negative loss for those charged with crimes is questionable. After all, the drafters of the Model Penal Code affirmatively sought to bolster mens rea protections beyond those provided by the common law, and their recommendation for universal

99. *See, e.g.*, Kenneth W. Simons, *Should the Model Penal Code's* Mens Rea *Provisions Be Amended?*, 1 OHIO ST. J. CRIM. L. 179, 179, 204 (2003) (observing how the Model Penal Code's mental state definitions and accompanying element analysis clarify and illuminate thinking about mens rea); Robinson & Grall, *supra* note 64, at 705–06 (same); Shen et al. *supra* note 88, at 1318 ("The [Model Penal Code] is now taught in virtually every law school[.]"); Kadish, *supra* note 88, at 521 (observing that the Model Penal Code is "the principal text in criminal law teaching" and "the point of departure for criminal law scholarship").

100. STUNTZ, *supra* note 18, at 260–62.

101. *Id.* at 261.

102. *Id.*

103. *Id.* at 262.

104. *See* Robert Weisberg, *Crime and Law: An American Tragedy*, 125 HARV. L. REV. 1425, 1444 (2012) (reviewing WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE (2011)) ("There is . . . a non sequitur in Stuntz's reading of *Morissette*, because the evil denounced in Justice Jackson's opinion—strict liability—is rejected in Stuntz's enemy, the [Model Penal Code]. Requiring knowledge in a theft case is hardly inconsistent with the [Model Penal Code's] use of arguably technical mens rea terms.").

culpable mental state requirements purports to do just that.[105] That said, it is also surely the case that replacing the common law's moralistic, open-ended standards with the narrower, more legalistic approach to mens rea reflected in the Model Penal Code has come at a cost.[106]

Those costs extend beyond the issues of prosecutorial administration highlighted by Stuntz. Arguably, the Model Penal Code's reconceptualization of mens rea has also had overlooked negative consequences on the creation of criminal policy. Although the precision of the PKRN framework has improved certain aspects of legislative deliberation, the "hypertechnical, cognitively based"[107] vision of mens rea it supplies may have detrimentally impacted others—for example, how lawmakers think about the moral salience of different mental states, and which mental states are sufficiently mitigating to merit exculpation. By untethering mens rea policy from any sense of why it matters, the PKRN approach has made it easier for legislatures to enact policies that authorize convictions for actors insufficiently blameworthy to warrant the condemnation of a criminal conviction.

I observed this dynamic while serving as the chief drafter of mens rea legislation for the only ground-up criminal code reform project to have

105. This point is discussed further *infra* notes 208–216 and accompanying text. For a more detailed exploration of the Model Penal Code's "strict liability abolition" agenda, see Serota, *supra* note 6, at 132–41.

106. *See* Joshua Kleinfeld, *Three Principles of Democratic Criminal Justice*, 111 NW. U. L. REV. 1455, 1464 n.22 (2017) (endorsing Stuntz's position that "common law *mens rea* standards, because of their moralistic and open-ended character, open up a necessary space for nontechnical argumentation about culpability and equity in criminal justice trials").

107. *See* Weisberg, *supra* note 104, at 1444.

occurred since the 1980s.[108] Working in this position for six years,[109] I witnessed how wholesale embrace of the Model Penal Code approach to mens rea has a tendency to divorce government deliberation from considerations of blameworthiness, while leading to a narrow, all-encompassing focus on PKRN which can make it difficult to appreciate the

108. Between 1962 and 1983, thirty-four jurisdictions adopted comprehensive criminal codes that "were influenced in some part by the Model Penal Code." Robinson & Dubber, *supra* note 85, at 326. Thereafter, in 1989, one additional jurisdiction, Tennessee, joined this group—a point often overlooked in the history of U.S. code reform. *See* State v. Williams, 38 S.W.3d 532, 535 (Tenn. 2001) (observing the state's adoption of a revised criminal code in 1989). Since then, no state has been able to successfully replace its chaotic collection of common law statutes with a modern, Model Penal Code-based code. *See* Robinson & Dubber, *supra* note 85, at 326. An individual state's efforts to do so is what I mean by "ground-up criminal code reform project."

Among the thirty-five jurisdictions that adopted modern, Model Penal Code-based codes, three have since attempted large-scale *re*-codification projects. They are Illinois, Kentucky, and, most recently, Delaware. *See* PAUL H. ROBINSON, FINAL REPORT OF THE ILLINOIS CRIMINAL CODE REWRITE AND REFORM COMMISSION (2003), https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1290&context=faculty_scholarship; PAUL H. ROBINSON, FINAL REPORT OF THE KENTUCKY PENAL CODE REVISION PROJECT (2003), https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1293&context=faculty_scholarship; PAUL H. ROBINSON, REPORT OF THE DELAWARE CRIMINAL LAW RECODIFICATION PROJECT (2017), https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=2748&context=faculty_scholarship; *see also* Michael T. Cahill, *Offense Grading and Multiple Liability: New Challenges for a Model Penal Code Second*, 1 OHIO ST. J. CRIM. L. 599 (2004) (discussing provisions from the Illinois and Kentucky projects, where Cahill also worked). To date, none of these comprehensive recodifications has been enacted—due to, among other reasons, opposition from law enforcement. *See, e.g.*, Paul H. Robinson, *The Rise and Fall and Resurrection of American Criminal Codes*, 53 U. LOUISVILLE L. REV. 173, 185 (2015) (discussing opposition from the Kentucky Attorney General's Office); Melissa Steele, *State Weighs Bills to Revamp Criminal System: Proposals Spark Response from Law Enforcement*, CAPE GAZETTE (Mar. 29, 2019), https://www.capegazette.com/article/state-weighs-bills-revamp-criminal-system/177438 (discussing similar opposition in Delaware).

109. Most of my legislative mens rea reform experience occurred in my capacity as Chief Counsel for Policy & Planning for the D.C. Criminal Code Reform Commission ("CCRC"), an independent agency in the District of Columbia, which prior to April 1, 2021, was focused on "develop[ing] comprehensive recommendations for the D.C. Council and Mayor on revision of District criminal statutes." *Mission*, D.C. CRIM. CODE REFORM COMM'N, https://ccrc.dc.gov/page/ccrc-mission (last visited Apr. 3, 2023). Final versions of the CCRC recommendations were required, by statute, to be submitted to the District of Columbia's Mayor and Council by March 31, 2021. D.C. CODE § 3-152 (2020). On November 15, 2022, the D.C. Council—the District of Columbia's local legislative body—unanimously approved something very close to these recommendations, in the form of the Revised Criminal Code Act of 2022 ("RCCA"). Martin Austermuhle, *D.C. Council Approves Sweeping Overhaul Of Criminal Code, Though Changes Won't Take Effect until 2025*, AM. UNIV. RADIO 88.5 (Nov. 15, 2022, 4:36 PM), https://dcist.com/story/22/11/15/dc-council-approves-major-overhaul-criminal-code/. In the ensuing months, however, the RCCA encountered significant opposition at the federal level, which ultimately stopped the new D.C. Code from being implemented. Martin Austermuhle, *How a D.C. Crime Bill Sparked a Political Firestorm and Ended Up Blocked By Congress*, AM. UNIV. RADIO 88.5 (Mar. 14, 2023, 2:58 PM), https://dcist.com/story/23/03/14/how-congress-blocked-dc-criminal-code-bill/. The future of the revised D.C. code is presently unclear.

possibility that something more than a culpable mental state requirement might be necessary to justify criminal liability.

By contrast, I also saw how conceptualizing mens rea in the broader and more moralistic terms reflected in the Guilty Minds approach can lead to more thoughtful discussions about criminal policy—and ultimately, more thoughtful criminal policies. The Guilty Minds approach accomplishes this, I believe, by making it easier for lawmakers to confront the big picture of mens rea: Why do mental states matter, which mental states are morally significant, and where might contemporary criminal policies fail to protect actors whose minds do not warrant the condemnation of a criminal conviction.? The next Section illustrates these dynamics through an analysis of the case of *Crossland v. United States*.

B. The Importance of Perspective: Crossland v. United States

On the afternoon of April 24, 2010, Terrance Crossland—a young Black man living in the District of Columbia—was mowing the lawn in front of his house while talking to his cousin.[110] While Crossland was mowing, two officers for the Metropolitan Police Department ("MPD") were engaged in an "aggressive high visibility patrol" in the area.[111] The officers, although aware that Crossland and his cousin were not doing anything unlawful, chose to detain the two men to procure information about an ongoing investigation—just one in a series of "unconstitutional" stops that MPD practiced at the time.[112]

Lacking "any right to go up and start searching" Crossland and his cousin, the officers aggressively "went up and seized them, told them to turn around, and started patting them down."[113] Crossland "'initially' complied, but quickly became 'agitated,' telling [the police] words to the effect of 'Fuck this shit. I'm tired of this.'"[114] What happened next is disputed, but it appears that Crossland intentionally elbowed one of the officers in the head, at which point all of the officers present beat Crossland with their fists and sprayed him with pepper spray.[115] During the scuffle, Crossland briefly resisted the attempts of one of the officers to place him in handcuffs.[116] On these facts, Crossland was convicted of two counts of Assault of a Police Officer

110. Crossland v. United States, 32 A.3d 1005 (D.C. 2011).
111. *Id.* at 1006.
112. *Id.* at 1009–10 (Schwelb, S.J., concurring).
113. *Id.* at 1009 (Schwelb, S.J., concurring).
114. *Id.* at 1006–07 (majority opinion).
115. *Id.*
116. *Id.* at 1008.

("APO")—one premised on the elbow, the other on the momentary resistance.[117]

The *Crossland* case leaves one with the sense that an injustice was done, but it can be difficult to identify its source. Punishing someone for conduct that would not have occurred but for illegal police misconduct is unsettling. But being harassed by the police does not excuse any and all forms of violence offered in response. For example, had Crossland strangled both of the officers to death, there is little doubt that Crossland ought to be held criminally responsible.

Perhaps, then, the sense of injustice is a product of the proportionality—and therefore justifiability—of Crossland's response under the circumstances. But that, too, does not seem correct. The intentional infliction of harm upon another is only justified when it is the right thing or a good thing to do.[118] And it is hard to label Crossland's actions as either. Crossland's response appears to have been retaliatory (instead of defensive), and the force he employed likely risked greater harm from the resulting confrontation than could have reasonably been prevented under the circumstances.[119] But even if Crossland's conduct was neither right nor good, his convictions may still seem unfair. To commit a crime, after all, is not just to do a bad thing. Instead, conduct becomes criminal only when there is a "concurrence of an evil-meaning mind with an evil-doing hand."[120] So, then, did Crossland possess the requisite evil-meaning mind?

The PKRN approach supplies an affirmative answer, based on the fact that Crossland acted with the most culpable mental state in the Model Penal Code's four-tier hierarchy: He *purposely* struck the first officer, and he *purposely* resisted the efforts of the second officer to arrest him.[121] No doubt, Crossland's decisions were influenced by a number of mitigating

117. *Id.* at 1006–07.

118. *See, e.g.*, Kadish, *supra* note 29, at 258 ("[The nature of a justification claim] is that I did nothing wrong even though I violated the prohibition."); DRESSLER, *supra* note 55, § 16.03 (defining justified conduct as conduct that violates a criminal prohibition yet society considers it to be "a good thing, or the right or sensible thing, or a permissible thing to do" under the circumstances).

119. For one thing, the District of Columbia's local legislative body, the D.C. Council, had placed a categorical bar on the use of force to resist an arrest, even if unlawful, with the hopes of deescalating the potential for violence for all parties. *In re* C.L.D., 739 A.2d 353, 355 (D.C. 1999) (noting that the policy behind D.C.'s APO statute is "to deescalate the potential for violence which exists whenever a police officer encounters an individual in the line of duty"). But also, there is an idea that "[i]f citizens who believe they are being wrongfully arrested will calmly submit to police officers who are detaining them, there will be fewer violent confrontations, and less people will get injured and killed." Craig Mackey, Note, Hudson v. Michigan *and the Ongoing Struggle for Accountability in Law Enforcement Institutions*, 6 ALB. GOV'T L. REV. 606, 631 (2013).

120. Morissette v. United States, 342 U.S. 246, 251 (1952).

121. This Article works with the version of the facts that the court accepted, which may not be an accurate representation of what actually happened.

circumstances; however, none of those circumstances would provide grounds for exculpation under the narrow body of excuse defenses that complement the PKRN approach.[122] For example, although Crossland seemed to be experiencing external coercion from the officer's unlawful and likely discriminatory search, the common law formulation of duress—most popular even in reform jurisdictions—is typically only available where death or severe bodily injury is threatened.[123] And while the Model Penal Code's approach to duress is materially broader, it is still limited to situations where someone was "so intimidated that he was unable to choose otherwise."[124] On the facts presented, Crossland's response appears to have been one of frustration and anger, not fear or intimidation.

While falling short of a conventional excuse defense, however, one may still question whether Crossland's state of mind was sufficiently blameworthy to justify two serious felony assault convictions.[125] For one thing, Crossland was experiencing *ethically appropriate* anger and frustration at the moment he employed force against the officers.[126] It is well-established that these kinds of emotions make it more difficult for people to think or act rationally.[127] And it is easy to see how Crossland's experiencing them could have impaired his ability to demonstrate absolute obeisance in the face of harassment perpetrated by officers from a department that had been systematically violating the constitutional rights of the people in his community.

Furthermore, there are other morally salient characteristics about Crossland's state of mind the court might have encountered had it taken the time to inquire further. For example, perhaps Crossland (or a close friend or relative) had experienced similar police abuses in the past—and observed,

122. *See supra* notes 89–91 and accompanying text.

123. *See, e.g.*, Robinson et al., *supra* note 89, at 88 (noting that the majority approach to duress requires the defendant to have been coerced "by what he reasonably believed was an unlawful threat of imminent death or severe bodily injury to himself or another"); DRESSLER, *supra* note 55, § 16.03.

124. MODEL PENAL CODE § 2.09 cmt. 2 at 373 (AM. L. INST. 1985) (focusing on a situation where an "actor makes a choice, but claims in his defense that he was so intimidated that he was unable to choose otherwise"); *see id.* at 376 ("The typical situation in which the section will be invoked is one in which the actor is told that unless he performs a particular criminal act a threatened harm will occur and he yields to the pressure of the threat, performing the forbidden act."); *see also* Robinson et al., *supra* note 89, at 88–89 (discussing relevant legal trends).

125. At the time, the District's APO statute provided a ten-year statutory maximum for commission of the offense. D.C. CODE § 22-405(a) (2007).

126. *See* Crossland v. United States, 32 A.3d 1005, 1008 (D.C. 2011) ("We discern no reason to doubt (and the government does not dispute) that Officer Baldwin's conduct—forcibly searching appellant when, as the officer acknowledged, appellant was doing nothing unlawful—violated appellant's Fourth Amendment right to be free from unreasonable searches and seizures.").

127. *See, e.g.*, Joshua Dressler, *Reflections on Excusing Wrongdoers: Moral Theory, New Excuses and the Model Penal Code*, 19 RUTGERS L.J. 671 (1988).

with overwhelming frustration, government officials continuously avoiding accountability. Perhaps the trauma associated with these experiences negatively impacted Crossland's relationships, employment prospects, or more general pursuit of a meaningful life. That these impairments fail to negate a purpose-to-cause harm or individually rise to the level of a narrow excuse defense does not make them morally immaterial.

Think of it this way: A criminal conviction and sentence say something morally condemnable about an offender—that they failed to care enough about the individual or public interests protected by the criminal law, and thus deserve the societal stigma and suffering associated with a formal pronouncement of blame.[128] Is this an accurate depiction of Crossland? Arguably it is not. Place anyone with sufficient concern for the safety and well-being of the police in Crossland's situation and they might have very well made the same choice—or a worse one. In which case it would be inappropriate to punish Crossland for failing to make a better choice.

This is just one example of how thinking about mens rea broadly, in terms of all mental characteristics that contribute to blameworthiness, can illuminate shortcomings in contemporary mens rea policy that solely focus on the PKRN approach can otherwise obscure. The problem, however, is that the Guilty Minds understanding of mens rea, while opening our eyes to underappreciated or overlooked problems, has historically been too vague and imprecise to provide concrete guidance or workable policy solutions. But what has been need not always be.

The next Part argues that there is an intelligible moral principle behind the Guilty Minds approach that is just waiting to be discovered. Synthesizing a wide range of experimental research, I find that the community's sense of psychological blameworthiness is mediated by four kinds of mental phenomena: motives, risk awareness, rationality, and volitional control. Working with these insights, I then construct a multi-dimensional model of culpability that offers a more refined conception of the Guilty Minds approach to mens rea.

II. THE PSYCHOLOGY OF GUILTY MINDS

What makes a mind guilty? There are many ways to approach the question, but let us begin with a simple observation: Much of what we care about in life is rooted in human minds and the decisions that arise from them.[129] For example, we celebrate good decisions, condemn bad decisions,

128. For further discussion of this idea, see *infra* Part III.

129. *See, e.g.*, Stephen J. Morse, *Criminal Law and Common Sense: An Essay on the Perils and Promise of Neuroscience*, 99 MARQ. L. REV. 39, 52 (2015) ("Virtually everything for which agents deserve to be praised, blamed, rewarded, or punished is the product of mental causation and, in principle, is responsive to reasons, including incentives.").

and withhold judgment for those—such as infants, pets, and those who have suffered severe brain injuries—whose ability to make decisions is limited or diminished.[130] This emphasis on the minds of others serves as a cornerstone of our emotional lives, the interpersonal relationships we form, and the societies we build together.[131] It also says something important about what it is to be a human: We are likely the only creatures who experience physical movements directed by a mind as having meaning or saying something important, both about the act and the actor.[132]

So, it should come as no surprise that when one person harms another person we take a keen interest in what was happening in the offending actor's mind.[133] For example, we instinctively gravitate toward things like intentions, motivations, awareness, and beliefs, all of which provide the most direct perspective on what this individual may have been thinking at the moment in question. And we also concern ourselves with broader mental phenomena from which we are able to determine whether the wrongdoing was perpetrated by someone with the ability to think and act morally (in contrast to, say, a small child or someone suffering from serious mental illness, who may not).

An illustration will be helpful. Imagine it is July 2020 and you are managing the entrance to a Costco in Phoenix that has just received its only shipment of medical supplies for the entire summer due to the supply chain disruptions resulting from a global pandemic. Although it is only 10:00 AM, it is already 117 degrees and sweltering. Right as you are about to let a young mother holding her small child, both of whom have been patiently waiting, through the sliding doors into the air conditioning, a towering man comes from out of nowhere, cuts the line, and jostles them, knocking both the mother and child over and onto the ground. Reflexively, and with great aggravation, you scowl at the man and exclaim: "What are you thinking?"

130. *See, e.g.*, Shoemaker, *supra* note 42, at 911–12 ("We have many emotional responsibility-responses to ourselves and others, including admiration, disdain, shame, pride, regret, disappointment, approval, anger, and gratitude (among many others). Much of the time, we may feel that this entire range of responses is available for the agents we come across, but for some agents we hesitate, feeling that only *some* of these responses are appropriate, whereas other responses would be inappropriate.")

131. *See generally* Strawson, *supra* note 42; WALLACE, *supra* note 42.

132. *See, e.g.*, Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1507 (2000) ("The expression of a mental state brings that state into the open, for oneself and potentially for others to recognize."); Christopher Eve Franklin, *Valuing Blame*, *in* BLAME: ITS NATURE AND NORMS, *supra* note 42, 207, 218–19 (observing that "burning a Picasso painting because of a dare carries a very different meaning from my burning a Picasso painting because of how cruelly it represents women"); Kimberly Kessler Ferzan, *Holistic Culpability*, 28 CARDOZO L. REV. 2523, 2533 (2007) ("Mental states are essential ingredients in meaning.").

133. *See, e.g.*, Ferzan, *supra* note 132, at 2532 ("[T]here is a constitutive relationship between the internal states of the actor and our assessment of moral blameworthiness.").

In posing this question you are unlikely to be interested in most of the things on this man's mind—for example, his plans for lunch or favorite movie. Rather, what you would likely want to know more about are those mental characteristics that are of moral relevance to the offensive act in question. For example, did the man *intend* to knock the mother over? Was he *aware* that the mother was holding a child? What was he *hoping* to achieve by jumping the line? And was his decision the product of a mind capable of *rational reflection* and *free choice*? Or was it distorted by the influence of psychotic thinking or a heat-induced emotional breakdown?

As these questions illustrate, the forms of mental phenomena that grab our attention are varied and complex—but their ethical import is often intuitive. A wide body of research in moral psychology indicates that we often possess an immediate feel for normative salience of different mental states, as well as a broader sense of responsibility that arises (or fails to arise) as we account for the minds of wrongdoers.[134] In what follows, I synthesize contemporary experimental and philosophical work to reveal four primary domains of interest in our evaluations of psychological blameworthiness.[135]

* * *

Question No. 1: *Why did the wrongdoer do it?* When confronted with wrongdoing, the first, and perhaps most powerful, domain of psychological interest people tend to gravitate towards is the nature of the wrongdoer's

134. For illuminating discussions about our shared capacity for rendering moral responsibility judgements and the influence that mental states have on those judgments, see, for example, JOHN M. DORIS & THE MORAL PSYCH. RSCH. GRP., THE MORAL PSYCHOLOGY HANDBOOK (2010); HOFFMAN, *supra* note 2; Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829 (2007); Norman J. Finkel, *Commonsense Justice, Culpability, and Punishment*, 28 HOFSTRA L. REV. 669 (2000) [hereinafter Finkel, *Commonsense Justice, Culpability, and Punishment*]; Norman J. Finkel, Marsha B. Liss & Virginia R. Moran, *Equal or Proportionate Justice for Accessories? Children's Pearls of Proportionate Wisdom*, 18 J. APPLIED DEVELOPMENTAL PSYCH. 229 (1997).

135. To be clear, this synthesis does not purport to capture the *only* mental characteristics relevant to psychological blameworthiness. *See, e.g.*, Garvey, *supra* note 33, at 556–57 (highlighting that "[o]ur reactive emotions" also seem to respond to an actor's awareness/ignorance of the illegality of their behavior, which is excluded from the discussion and ensuing culpability model developed in this Part). There are also other ways to conceptualize the psychological domains discussed in this Part. *See, e.g.*, Kimberly Kessler Ferzan, *Plotting Premeditation's Demise*, 75 L. & CONTEMP. PROBS. 83, 93–99 (2012) (proposing a five-prong analysis of culpable choice to be used to identify the most culpable killings). The domains discussed below are thus an attempt to capture those mental characteristics that are most integral to blame, as established in relevant moral psychology research, in a manner that is most susceptible to being evaluated by legal decisionmakers. *See infra* IV.B (translating the Guilty Minds culpability model into an affirmative insufficient blameworthiness defense).

motivations.[136] More specifically, we want to understand an actor's reasons for engaging in wrongful conduct—and have morally distinct responses to their wrongdoing contingent upon the nature of those reasons.[137] All else being equal, the worse an actor's reasons for engaging in wrongdoing are, the more blameworthy people tend to perceive a wrongdoer to be.[138]

To illustrate, consider the following variations in the motivation behind the line-cutter's jostle:

(1) The line-cutter acted to help the woman and her child avoid what he mistakenly perceived to be a car about to sideswipe them;

136. For philosophical work emphasizing the importance of motivations to moral and criminal responsibility, see, for example, Ferzan, *supra* note 135, 104–05; ALEXANDER & FERZAN, *supra* note 29, at 23–31, 93–103; Guyora Binder, *Making the Best of Felony Murder*, 91 B.U. L. REV. 403, 434–37 (2011); SCANLON, *supra* note 42, at 124–25; Carissa Byrne Hessick, *Motive's Role in Criminal Punishment*, 80 S. CAL. L. REV. 89 (2006). For empirical work suggesting that the public holds similar views, see, for example, PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW (1995) (Studies 5, 6, 8, 9, and 16); Bertram F. Malle, Steve Guglielmo & Andrew E. Monroe, *A Theory of Blame*, 25 PSYCH. INQUIRY 147, 175 (2014) ("[J]ustification is a continuous value, varying with the degree of credibility and cultural acceptability of the provided reasons and with the extremity of the norm violation." (citing Dov Cohen & Richard E. Nesbitt, *Self-Protection and the Culture of Honor: Explaining Southern Violence*, 20 PERSONALITY & SOC. PSYCH. BULL. 551 (1994))); Janice Nadler & Mary-Hunter McDonnell, *Moral Character, Motive, and the Psychology of Blame*, 97 CORNELL L. REV. 255, 280–92 (2012); Glenn D. Reeder et al., *Inferences About the Morality of an Aggressor: The Role of Perceived Motive*, 83 J. PERSONALITY & SOC. PSYCH. 789 (2002); Robinson & Darley, *supra* note 91; Norman J. Finkel & Jennifer L. Groscup, *Crime Prototypes, Objective Versus Subjective Culpability, and a Commonsense Balance*, 21 LAW & HUM. BEHAV. 209 (1997); Martin F. Kaplan, *Judgments of Murder Mysteries as a Means of Studying Juror Cognition*, 10 BASIC & APPLIED SOC. PSYCH. 299 (1989); *see also* PAUL H. ROBINSON, INTUITIONS OF JUSTICE AND THE UTILITY OF DESERT 313–14 (2013) (finding across a number of studies involving accomplice liability that people assign more liability and punishment for purpose, as opposed to knowledge or recklessness, as to facilitating criminal conduct).

It is important to acknowledge that criminal law jurists and scholars often claim that "motivation" is immaterial to liability and punishment. *See, e.g.*, Elaine M. Chiu, *The Challenge of Motive in the Criminal Law*, 8 BUFF. CRIM. L. REV. 653, 656 (2005) ("Generations of scholars of the criminal law have learned that motive is irrelevant in the criminal law."); Douglas N. Husak, *Motive and Criminal Liability*, 8 CRIM. JUST. ETHICS 3, 3 (1989) ("This thesis is endorsed, sometimes with minor qualifications, by almost all leading criminal theorists."). But this claim is at best misleading: From justification defenses, to bias crimes, to the gross deviation standards governing the common definitions of recklessness and negligence, to sentencing decisions, the criminal law accounts for motivations in diverse ways. *See, e.g.*, Hessick, *supra*. And this is to say nothing of the informal ways that motive likely influences the decisions of courts, prosecutors, and juries. *See, e.g.*, Anna Roberts, *Dismissals as Justice*, 69 ALA. L. REV. 327, 376 (2017) (providing illustrations in the context of *de minimis* dismissals). This should not be surprising, however, given that a prohibition on considering motive "violates widely held intuitions." Binder, *supra* note 136, at 426.

137. Malle et al., *supra* note 136, at 154 ("When moral perceivers regard the negative event in question as intentional . . . they consider the agent's particular reasons for acting. . . . Considering an agent's reasons is an intrinsic part of the moral perception of intentional actions because these reasons determine the *meaning* of the action . . . .").

138. *See, e.g.*, sources cited *supra* note 136.

(2) The line-cutter acted to ensure that the line-cutter and his family would have necessary medical supplies for the summer.
(3) The line-cutter acted to send a message of racial inferiority to the young woman and her child.

Because these motivations are ranked according to their moral superiority, existing research indicates that people will tend to view the actors in these three scenarios as manifesting increasingly greater degrees of psychological blameworthiness.[139]

Question No. 2: *Whether and to what extent was the wrongdoer aware of the risks involved?* When confronted with wrongdoing, the second domain of psychological interest people tend to gravitate towards is the nature of an actor's awareness of the risks associated with their wrongful conduct.[140] More specifically, our interest in the awareness accompanying wrongdoing seems to track two different variables: (1) the level of awareness possessed by the wrongdoer (e.g., awareness that harm was *possible* versus awareness that harm was *nearly certain* to occur); and (2) the gravity of the harm of which the actor was aware (e.g., awareness that a *serious injury* was possible

139. *See, e.g.*, sources cited *supra* note 136.

140. For philosophical work emphasizing the importance of risk awareness to moral and criminal responsibility, see, for example ALEXANDER & FERZAN, *supra* note 29, at 23–31; Alexander F. Sarch, *Willful Ignorance, Culpability, and the Criminal Law*, 88 ST. JOHN'S L. REV. 1023, 1062–65 (2014); Gideon Yaffe, *The Point of Mens Rea: The Case of Willful Ignorance*, 12 CRIM. L. & PHIL. 19, 28–32 (2018). For empirical work suggesting that the public holds similar views, see, for example, ROBINSON & DARLEY, *supra* note 136 (Studies 5, 6, 8, 9, and 16); ROBINSON, *supra* note 136, at 305–06, 313–14, 327; David A. Lagnado & Shelley Channon, *Judgments of Cause and Blame: The Effects of Intentionality and Foreseeability*, 108 COGNITION 754 (2008); Lara Kirfel & Ivar Rodríguez Hannikainen, *Why Blame the Ostrich? Understanding Culpability for Willful Ignorance*, *in* ADVANCES IN EXPERIMENTAL PHILOSOPHY OF LAW (Stefan Magen & Karolina Prochownik eds.) (forthcoming May 2023), https://psyarxiv.com/kswtu/; Malle et al., *supra* note 136, at 155 ("Agents who cause a norm-violating event that they foresaw . . . receive more blame than agents who cause a norm-violating event that they did not and could not foresee (holding physical capacity constant).").

It is important to note that two recent studies call into question (1) whether lay jurors can reliably make fine-grained distinctions as to the extent of an actor's conscious risk awareness and (2) whether material variances in conscious risk awareness, even when accurately identified by lay jurors, are viewed as increasing an actor's blameworthiness. Matthew R. Ginther et al., *The Language of Mens Rea*, 67 VAND. L. REV. 1327, 1339–61 (2014); *id.* at 1360 (finding "conflation of K and R punishment because subjects do not see a clear moral distinction between the K and R mental states, at least as it concerns the result element of offenses"); Shen et al., *supra* note 88, at 1326–55; *see also* Francis X. Shen, *Minority Mens Rea: Racial Bias and Criminal Mental States*, 68 HASTINGS L.J. 1007, 1017–42 (2017) (conducting a study reaffirming that lay jurors struggle to distinguish between material variances in an actor's conscious risk awareness but also finding that culpable mental state evaluations of this nature may be resistant to implicit racial bias). For a fascinating study employing neuroimaging and machine-learning techniques to explore the neural correlates of variances in conscious risk awareness, see Iris Vilares et al., *Predicting the Knowledge—Recklessness Distinction in the Human Brain*, 114 PROC. NAT'L ACAD. SCIS. 3222 (2017); *see also* Owen D. Jones, Read Montague & Gideon Yaffe, Essay, *Detecting Mens Rea in the Brain*, 169 U. PA. L. REV. 1 (2020) (discussing this study).

versus awareness that a *minor injury* was possible).[141] All else being equal, the greater one's level of culpable awareness and the graver the harm of which the wrongdoer is aware, the more blameworthy people tend to perceive the wrongdoer to be.[142]

Consider, for example, the following variations in states of awareness accompanying the line-cutter's jostle:

(1) The line-cutter was *completely unaware* that his conduct posed *any risk of injury* to either the mother or child.
(2) The line-cutter was aware of a *possibility* that his conduct would cause *a minimal injury* to both the mother and child.
(3) The line-cutter was aware that his conduct was *practically certain* to cause *serious injury* to both the mother and child.

Because each of these scenarios reflects comparatively greater degrees of culpable awareness, existing research indicates that people will tend to view them as revealing increasingly greater degrees of psychological blameworthiness.[143]

Question No. 3: *Whether and to what extent was the wrongdoer able to engage in moral reasoning?* When confronted with wrongdoing, the third domain of psychological interest people tend to gravitate towards is a wrongdoer's capacity to engage in moral reasoning.[144] More specifically, our

141. There are many other dimensions to risk awareness, such as, for example, the difference between *knowing* something to be true and *believing* something to be true, which does not in fact turn out to be true. *See, e.g.*, Alan C. Michaels, *Acceptance: The Missing Mental State*, 71 S. CAL. L. REV. 953, 1032 n.330 (1998). These differences are set aside for purposes of the present discussion.

142. *See, e.g.*, sources cited *supra* note 140.

143. *See, e.g.*, sources cited *supra* note 140.

144. For philosophical work emphasizing the relationship between blameworthiness and the extent to which an agent possesses the rational capacities necessary to distinguish right from wrong, see, for example, R.A. DUFF, ANSWERING FOR CRIME: RESPONSIBILITY AND LIABILITY IN THE CRIMINAL LAW 39–40 (2007); David O. Brink & Dana K. Nelkin, *Fairness and the Architecture of Responsibility*, *in* 1 OXFORD STUDIES IN AGENCY AND RESPONSIBILITY 284 (David Shoemaker ed., 2013); Stephen J. Morse, *Excusing and the New Excuse Defenses: A Legal and Conceptual Review*, 23 CRIME & JUST. 329, 340 (1998); ALEXANDER & FERZAN, *supra* note 29 at 155 ("Rationality is the cornerstone of responsible agency. If an actor cannot comprehend or respond to norms, then it cannot be said that laws or morality are properly addressed to the actor." (footnote omitted)); *see also* Stephen J. Morse, *Reason, Results, and Criminal Responsibility*, 2004 U. ILL. L. REV. 363, 440 ("[T]he capacity for rationality is a congeries of skills, including the ability to perceive accurately, to reason instrumentally according to a minimally coherent preference-ordering, and to appreciate the significance of reasons and their connection to our actions."). For empirical work suggesting that the public holds similar views, see, for example, ROBINSON & DARLEY, *supra* note 136 (Studies 12 & 13); ROBINSON, *supra* note 136, at 342 ("Perpetrators who are judged to be suffering from a high degree of [cognitive or control] dysfunction . . . are normally not assigned criminal liability."); *id.* at 348 ("[B]oth cognitive and control dysfunction appear to support a defense [for involuntary intoxication]."); Lane Kirkland Gillespie et al., *Examining the Impact of Proximate Culpability Mitigation in Capital Punishment Sentencing Recommendations: The Influence of Mental Health Mitigators*, 39 AM. J. CRIM. JUST. 698, 707–08, 710 (2014) (finding that, in capital cases, a jury's

interest in a wrongdoer's capacity for moral reasoning seems to track whether and to what extent cognitive or situational factors beyond the control of the actor hindered their ability to distinguish right from wrong. All else being equal, the greater the influence of psychological impairments to a wrongdoer's capacities for moral reasoning, the less blameworthy people tend to perceive the wrongdoer to be.[145]

Consider, for example, the following variations in impairment to the line-cutter's capacity for moral reasoning:

(1) The line-cutter was experiencing extreme psychotic delusions, which made it impossible for him to recognize the immorality of injuring the mother and child.
(2) The line-cutter was experiencing low-level psychotic delusions, which made it meaningfully more difficult for him to recognize the immorality of injuring the mother and child.
(3) The line-cutter was not experiencing any diminishment in his ability to distinguish right from wrong.[146]

Because each of these scenarios reflects comparatively lesser degrees of impairment to one's capacities for moral reasoning, existing research indicates that people will tend to view them as revealing increasingly greater degrees of psychological blameworthiness.[147]

Question No. 4: *Whether and to what extent was the wrongdoer able to control their conduct?* When confronted with wrongdoing, the fourth domain of psychological interest people tend to gravitate towards is volitional control.[148] More specifically, our interest in a wrongdoer's ability to exercise

---

acceptance of a defendant's reduced capacity to appreciate criminality of conduct lowered the probability of a death sentence recommendation by 18.31%); Stephen P. Garvey, *Aggravation and Mitigation in Capital Cases: What Do Jurors Think?*, 98 COLUM. L. REV. 1538, 1564–65 (1998) (finding that "a majority of jurors would be at least slightly less likely to vote for death if the defendant had a history of mental illness"); James R.P. Ogloff, *A Comparison of Insanity Defense Standards on Juror Decision Making*, 15 LAW & HUM. BEHAV. 509, 526–28 (1991); Norman J. Finkel et al., *Insanity Defenses: From the Jurors' Perspective*, 9 LAW & PSYCH. REV. 77, 81–84 (1985).

145. *See, e.g.*, sources cited *supra* note 144.

146. This example analyzes the relationship between blame and mental illness in terms of "mental illness's effect on the actor's ability to appreciate the wrongness of imposing the risks for the reasons he has for doing so." ALEXANDER & FERZAN, *supra* note 29, at 157. However, mental illness can also have an "effect on the actor's ability to refrain from acting culpably—that is, by its effect on his volitional control." *Id.* Question No. 4, below, explores this relationship but as it relates to involuntary intoxication, instead of mental illness.

147. *See, e.g.*, sources cited *supra* note 144.

148. For philosophical work emphasizing the relationship between blameworthiness and the extent to which an agent possesses the volitional capacities necessary to conform their conduct to their moral judgments, see, for example, Brink & Nelkin, *supra* note 144; HART, *supra* note 29, at 23; Paul H. Robinson, *A System of Excuses: How Criminal Law's Excuse Defenses Do, and Don't, Work Together to Exculpate Blameless (and Only Blameless) Offenders*, 42 TEX. TECH L. REV. 259 (2009); Joshua Dressler, *Some Very Modest Reflections on Excusing Criminal Wrongdoers*, 42 TEX.

this form of behavioral regulation tracks whether and to what extent cognitive or situational factors beyond the control of the actor hindered their ability to conform their conduct to what morality requires.[149] All else being equal, the greater the influence of relevant volitional impairments from which a wrongdoer suffered, the less blameworthy people tend to perceive the wrongdoer to be.[150]

Consider, for example, the following variations in impairment to the line-cutter's capacity to control his conduct (although the line-cutter recognized what he did to be wrong):

(1) The line-cutter was completely unable to control his conduct because of an unexpected side effect of a prescription medication, which the line-cutter's negligent doctor had failed to warn him about.
(2) The line-cutter was substantially unable to control his conduct because of an unexpected side effect of a prescription medication, which the line-cutter's negligent doctor had failed to warn him about.
(3) The line-cutter was fully capable of controlling his conduct.

Because each of these scenarios reflects increasingly lesser levels of volitional impairment, existing research indicates people will tend to view them as manifesting increasingly greater degrees of psychological blameworthiness.[151]

* * *

TECH L. REV. 247, 256–57 (2009). Matthew Talbert, *Implanted Desires, Self-Formation and Blame*, 3 J. ETHICS & SOC. PHIL. Aug. 2009, at 11–12 ("[T]he question of whether it is reasonable to blame Beth will be best answered by inquiring into whether she is capable of governing her behavior according to internal values and judgments so that her behavior expresses interpersonally significant attitudes."). For empirical work suggesting that the public holds similar views, see, for example, ROBINSON & DARLEY, *supra* note 136 (Studies 12 & 13); ROBINSON, *supra* note 136, at 342, 348; Garvey, *supra* note 144, at 1564–65; Glenn D. Reeder et al., *Impressions of Milgram's Obedient Teachers: Situational Cues Inform Inferences About Motives and Traits*, 95 J. PERSONALITY & SOC. PSYCH. 1 (2008) (finding that research participants were able to identify the impact of situational forces on another's actions and attributed less blame when there was greater coercion); Robert L. Woolfolk et al., *Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility*, 100 COGNITION 283 (2006) (finding that study participants who were given hypothetical stories were more likely to deem actors morally responsible the less external impairment the actors experienced); Malle et al., *supra* note 136, at 176 ("[U]nder extreme social pressure or duress . . . the community acknowledges that the agent behaved like any reasonable person would and therefore reduces blame.").

149. It bears notice that the source of an actor's loss of control "may be either external, such as a gun at one's head, or internal, such as an alleged lack of control capacity produced by mental disorder." Stephen J. Morse, *Severe Environmental Deprivation (aka RSB): A Tragedy, Not a Defense*, 2 ALA. C.R. & C.L. L. REV. 147, 148 (2011).

150. *See, e.g.*, sources cited *supra* note 148.

151. *See. e.g.*, sources cited *supra* note 148.

That people tend to agree on the extent to which individual mental characteristics contribute to blameworthiness is notable. But even more impressive is just how effortlessly people seem to be able to synthesize disparate psychological phenomena into shared determinations of moral and criminal responsibility. To illustrate, consider three possible psychological profiles of the line-cutter that aggregate different combinations of the domain-specific mental states described above.

(1) The line-cutter was motivated by a desire to send a message of racial inferiority to both the mother and the community. At the time of the act, he believed that he would almost surely cause serious bodily injury to both the mother and child. In addition, the line-cutter was not suffering from any cognitive or volitional impairments at the time of the act; rather, he was fully capable of engaging in rational reflection and was also in full control of his conduct.

(2) The line-cutter was motivated by a desire to secure necessary medical supplies for his family. At the time of the act, he believed that his conduct posed a small risk of minimal injury to both the mother and child. In addition, the line-cutter was suffering from a moderate mental illness at the time of the act, which made it meaningfully more difficult for him to distinguish right from wrong.

(3) The line-cutter was motivated by a desire to save the woman and child from what he mistakenly perceived to be a car about to run them over. At the time of the act, he perceived no risk of harm to either the woman or the mother from the jostling—and even if he did, it would not have mattered. Because earlier in the day, the line-cutter took a prescription medication (for the first time), at the behest of his negligent doctor, which made it impossible for him to control his conduct.

Reading through these examples, one may find it easy (or at least easier than otherwise assumed) to answer a couple of questions. The first is binary (yes/no): Are each of these actors blameworthy? And the second is comparative: How does the blameworthiness of each actor compare with the others on a single moral continuum? That is to say, among the three scenarios, which actor is: (i) most blameworthy; (ii) least blameworthy; and (iii) who falls in between? Existing research indicates that many people will converge on their assessments—finding, for example, that the first actor clearly is blameworthy, that the third actor clearly is not, and that each of the three

profiles rests on a continuum of descending psychological blameworthiness.[152]

Convergence in the community's sense of psychological blameworthiness is impressive, particularly given the diversity of moral perspectives that exist in a place like the United States. But that convergence is also far from complete—so it is important to be clear about limitations. What I've described here are only *patterns* of moral judgment reflected in relevant psychological, legal, and philosophical literature. And these patterns are far from monolithic, while moral decision-making by lay jurors is influenced by a diversity of factors.[153] Experimental research therefore also reveals instances where community sentiment appears to be split over the relevance of mens rea, or converge on its irrelevance altogether.

So, for example, in some situations, people's moral responsibility judgments appear to focus on the gravity of the harm caused or threatened by a wrongdoer, without regard to their accompanying state of mind.[154] And in other situations, people's moral responsibility judgments seem to be significantly influenced by normatively irrelevant aspects of the person being judged, which range from the arbitrary to the reprehensible.[155] And yet, the

152. *See, e.g.*, ROBINSON & DARLEY, *supra* note 136, at 208–09; ROBINSON, *supra* note 136; *see also* Norman J. Finkel & Jennifer L. Groscup, *When Mistakes Happen: Commonsense Rules of Culpability*, 3 PSYCH. PUB. POL'Y & L. 65, 120 (1997) ("Commonsense justice ends with fine-grain culpability distinctions, a gradation.").

153. For a good empirical accounting of these diverse sources, see NORMAN J. FINKEL, COMMONSENSE JUSTICE: JURORS' NOTIONS OF THE LAW (1995); *see also* Finkel, *Commonsense Justice, Culpability, and Punishment*, *supra* note 134, at 701 (observing that "jurors are likely to flavor, combine, and cook them in more subjective and psychological ways, throwing in past experiences, intuitions, sentiments, biases, heuristics, construals, and prototypes, as they wok and roll").

154. For discussion of harm's role in community assessments of blameworthiness, see *infra* notes 241–242 and accompanying text. And for studies indicating public support for certain forms of strict liability, see ROBINSON & DARLEY, *supra* note 136, at 88–89 tbl.4.1 (finding that the vast majority of respondents would impose criminal liability in situations involving accidental damage to property, although the individual acted non-negligently); Carly Giffin & Tania Lombrozo, *Wrong or Merely Prohibited: Special Treatment of Strict Liability in Intuitive Moral Judgment*, 40 LAW & HUM. BEHAV. 707 (2016); Joseph Sanders et al., *Must Torts Be Wrongs? An Empirical Perspective*, 49 WAKE FOREST L. REV. 1 (2014). But note: researchers have found that, "when the head-to-head evidence" between the influence of intent and harm on moral judgment "is examined, th[e] hegemony goes to intent, as it remains the starting point, and often the final point, in commonsense justice's culpability analysis." Finkel & Groscup, *supra* note 152, at 117.

155. For example, as Vera Bergelson observes, "[p]ublic views on the allocation of responsibility for rape are well known for their unfairness to the victim," such that reliance on community sentiment might support "a 'mini-skirt' defense to the crime of rape." Vera Bergelson, *Victims and Perpetrators: An Argument for Comparative Liability in Criminal Law*, 8 BUFF. CRIM. L. REV. 385, 428–30 (2005). Other empirical work finds the influence of "blatant biases and base sentiments." Finkel, *Commonsense Justice, Culpability, and Punishment*, *supra* note 134, at 702; *see, e.g.*, Avani Mehta Sood, *Attempted Justice: Misunderstanding and Bias in Psychological Constructions of Criminal Attempt*, 71 STAN. L. REV. 593, 639 (2019) ("Lay decisionmakers applying the proximity standard in an attempted terrorism case were significantly more likely to

fact remains that we can recognize these as *departures* precisely because there exists a coherent framework of psychological blameworthiness that is consistent with our basic moral commitments as well as what we think, say, and do about blame much of the time.

That framework, to briefly recap, understands the moral salience of the minds of wrongdoers to exist on a continuum mediated by four central domains of psychological inquiry: motivation, awareness, rationality, and volitional control.[156] The interaction between the mental characteristics that comprise these domains is principally what yields the community's perspective on psychological blameworthiness. And that perspective, in turn, functions in one of two ways: (1) on a threshold/binary level (blameworthy vs. not blameworthy); and (2) on a gradational level, which admits of greater and lesser degrees of blameworthiness and allows for comparative assessments of culpability to be made between different types of actors.

This more refined understanding of psychological blameworthiness captures the essence of the common law's vague and moralistic vision of mens rea while offering greater precision and specificity. In so doing, this understanding supplies the Guilty Minds approach with the conceptual architecture that it has historically been missing, and which has thus far precluded lawmakers from fully realizing the benefits of that approach for the criminal law.

Part IV of this Article will deploy the Guilty Minds culpability model as the basis for a novel doctrinal solution to structural flaws in PKRN mens rea policies that authorize convictions for insufficiently blameworthy actors. But before discussing *how* to limit criminal liability to psychologically blameworthy actors, it is necessary to address the logically prior question of *why* criminal liability should be so limited? That is the focus of Part III.

---

construe the thoughts and actions of the defendant as criminal when his name suggested that he was Muslim—even when the legally relevant evidence in the case skewed toward innocence and had nothing to do with Islam."); HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY 164–66 (1966) (postulating a "liberation hypothesis": that personal sentiments will be more likely to color jurors' judgments when the legally relevant evidence in a case is ambiguous, because "doubts about the evidence free the jury to follow sentiment"); *see also, e.g.*, Robert J. MacCoun, *Biases in the Interpretation and Use of Research Results*, 49 ANN. REV. PSYCH. 259, 272–73 (1998) (discussing Kalven and Zeisel's liberation hypothesis); Robert J. MacCoun, *The Emergence of Extralegal Bias During Jury Deliberation*, 17 CRIM. JUST. & BEHAV. 303, 306 (1990) (same). *But see* Shen, *supra* note 143, at 1030 ("Do scenario protagonists named Jamal and Lakisha receive more culpable mental state assessments than counterparts named John and Emily? The results suggest that the answer to this question is *no*.").

156. As discussed *supra* note 135, this framework only purports to capture mental characteristics that are most salient to blameworthiness based on existing research in a way that is susceptible to being effectively administered by legal decisionmakers. *See infra* Part IV (translating this culpability model into an affirmative insufficient blameworthiness defense).

III. THE MORALITY OF GUILTY MINDS

While psychological blameworthiness may be an integral part of the community's moral responsibility judgments, the community's moral responsibility judgments are not necessarily an integral part of criminal policy. The relationship between criminal policymaking and public opinion is complicated by a range of political phenomena, including legislative ignorance and interest group pressure.[157] And these complications are particularly pronounced when it comes to mens rea, which is a challenging topic on which public opinion is not well-known to those involved in the creation of criminal policy and for which key legislative constituencies (e.g., prosecutors and law enforcement) hold positions more punitive than those held by the public.[158]

Under these circumstances, academic research has an important role to play in teaching lawmakers and reformers why mens rea matters to criminal policy. Historically, however, mens rea scholarship has failed to serve this critical function.[159] I have sought to address this problem in a pair of recent projects that articulate the empirical case—understood in terms of public safety, decarceration, and political expediency—in support of universal culpable mental state requirements.[160] This Part complements that work by articulating the moral philosophical case—understood in terms of fairness and general societal welfare—in support of viewing the guilty mind, understood in terms of psychological blameworthiness, as a critical limit on criminal convictions.

The key to that argument, as explained in Section III.A, lies in recognizing what criminal convictions *are*—namely, formalized public expressions of blame—and what this particular class of expressions *says*—namely, that someone failed to care enough about valuable individual and societal interests. By understanding the social meaning of blame and the moral significance of mental states, as explained in Section III.B, we come to see why those lacking a guilty mind should be firmly excluded from the scope of liability contained within U.S. criminal codes.

157. For an illuminating discussion of this idea, as well as penal populism's role in criminal policymaking more generally, see Paul H. Robinson & Jonathan C. Wilt, *Undemocratic Crimes*, 2022 U. ILL. L. REV. 485, as well as the sources cited *supra* in notes 16 and 18.

158. *See* Robinson & Wilt, *supra* note 157, at 504–06, 508–10.

159. *See* Mizel, Serota, Cantor & Russell-Fritch, *supra* note 23, at 292 (discussing the shortcomings of contemporary mens rea scholarship). For two notable exceptions that provide uniquely accessible scholarly treatments of mens rea policy, see Levenson, *supra* note 8, and Kadish, *supra* note 29.

160. *See generally* Serota, *supra* note 6; Mizel, Serota, Cantor & Russell-Fritch, *supra* note 23.

A. Mental States and the Social Meaning of Blame

Why should criminal liability be limited to those who act with guilty minds? The best place to start in addressing this question is with a more careful examination of blame, which is arguably the criminal law's most important characteristic and defining feature. Part I of this Article generally considered what blame *is*, namely, a kind of "moral criticism" directed at someone for violating a community norm.[161] But another critical dimension to appreciate is what blame *does*, namely, it expresses "a judgment of an actor's values."[162] As Peter Westen writes:

> Morally, blame is indignation on the part of an agent, *A*, toward another agent, *B*, for *B*'s causing a morally wrongful state of affairs, *C*—indignation being reproach by *A* toward *B* for the latter's motivation in causing *C* and, specifically, for his causing *C* out of wrongful disregard for the legitimate interests of others . . . .[163]

The basic idea Westen is conveying—both here and in other work[164]—is that our blaming practices revolve around assessing (and ultimately, expressing) the level of concern (or lack thereof) wrongdoers have for the interests of other people. To put the point plainly, when we *publicly* blame someone for *moral* wrongdoing[165] we are in effect communicating something like: "Hey, you messed up. It is not just that you did something wrong, but in choosing to do the wrong thing, you revealed that you don't actually care enough about the interests of other people. And that's immoral!"

At first blush, the idea that moral blame is fundamentally about publicly condemning those who manifest insufficient concern for individual or societal interests[166] may sound ethereal. But on reflection, it is intuitive. Caring about human interests is indelibly connected to a basic tenet of our

161. Kadish, *supra* note 29, at 264.

162. Westen, *supra* note 46, at 151.

163. Peter Westen, Lex Mitior*: Converse of* Ex Post Facto *and Window into Criminal Desert*, 18 NEW CRIM. L. REV. 167, 187–88 (2015) (emphasis added).

164. *See id.*; Westen, *supra* note 29; Westen, *supra* note 46.

165. There are, as Westen observes, two different interpretations of moral blame: "Subjectively, moral blame is emotion or sentiment of indignation on *A*'s part toward *B*, whether *A* expresses it or not. Objectively, moral blame is a public expression of indignation by *A* toward *B*, whether *A* personally feels the indignation or not." Westen, *supra* note 163, at 188 (emphasis added). This Article focuses on the latter, publicly expressed variety of moral blame.

166. Note that a person may violate individual interests, such as, for example, where *D* punches *V*, whether or not punching is legally prohibited. Or a person may violate societal interests, such as, for example, where *D* sells *X* a form of contraband, where the transfer of contraband is legally prohibited. *See also* YAFFE, *supra* note 29, at 78 ("[I]t is perfectly possible to legally prohibit acts in virtue of features that have nothing whatsoever to do with the relevance of those acts to other people."). Arguably, all legal interests (i.e., those protected by legal norms) are societal interests, but not all societal interests are legal interests.

social morality: Most of us believe, or at least live our lives under the assumption that, we all *owe* one another some degree of concern.

Just how far this obligation extends varies across relationships, contingent upon the level intimacy and dependency involved.[167] Most would presumably accept, for example, that a parent owes their child a greater level of concern than, say, a random person on the street. And yet, even to that random person, it is generally uncontroversial to think that some minimum level of concern is owed—say, for example, to refrain from knocking them over, hitting them with a car, or imposing other serious risks to their safety or well-being in the absence of strong countervailing reasons.[168] So it is not surprising that, when people fall short of their obligation to minimally concern themselves with the safety and well-being of others, it is established societal practice to publicly call out the violation of the norm.[169] That is blame.

Where do minds fit into this picture? The mind is, simply put, where one's concern (or again, lack thereof) for human interests registers.[170] To

167. *See, e.g.*, SCANLON, *supra* note 42, at 127–28 (discussing the Strawsonian understanding that "[d]ifferent relationships involve different [moral] standards.").

168. *See, e.g.*, *id.* at 140 ("[Morality] requires us to take care not to behave in ways that will harm those to whom we stand in this relation, to help them when we can easily do so, not to lie to them or mislead them, and so on").

169. There exist multiple literatures that emphasize the relationship between insufficient concern and blame. For important contributions from the criminal law literature, see, for example, Simons, *supra* note 96; YAFFE, *supra* note 29; Yaffe, *supra* note 26; Westen, *supra* note 29; ALEXANDER & FERZAN, *supra* note 29; Garvey, *supra* note 33; Alexander Sarch, *Who Cares What You Think? Criminal Culpability and the Irrelevance of Unmanifested Mental State*s, 36 L. & PHIL. 707 (2017). *See also* Samuel H. Pillsbury, *The Meaning of Deserved Punishment: An Essay on Choice, Character, and Responsibility*, 67 IND. L.J. 719 (1992); Michaels, *supra* note 141; Joshua Kleinfeld, *Why the Mind Matters in Criminal Law*, 53 ARIZ. ST. L.J. 539 (2021). For important contributions from the realm moral philosophy, see, for example, JOHN MARTIN FISCHER & MARK RAVIZZA, RESPONSIBILITY AND CONTROL: A THEORY OF MORAL RESPONSIBILITY (1998); Angela M. Smith, *Responsibility for Attitudes: Activity and Passivity in Mental Life*, 115 ETHICS 236 (2005); SCANLON, *supra* note 42; Pamela Hieronymi, *Controlling Attitudes*, 87 PAC. PHIL. Q. 45 (2006). Within this body of literature, there are important disagreements. For example, criminal law theorists dispute whether negligent inadvertence can manifest the kind of insufficient concern necessary to ground criminal liability and punishment. Kenneth W. Simons, Review, *Retributivism Refined—or Run Amok?*, 77 U. CHI. L. REV. 551, 566–68 (2010) (reviewing ALEXANDER & FERZAN, *supra* note 29) (discussing different ways the insufficient concern principle might be construed, and the implications for negligence liability); *see also* Douglas Husak, *Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting*, 5 CRIM. L. & PHIL. 199, 199 (2011) ("Should criminal liability ever be imposed for negligent conduct? Commentators disagree radically.").

170. *See, e.g.*, ALEXANDER & FERZAN, *supra* note 29, at 6 (observing that blameworthiness is rooted in the "defendant's decision to violate society's norms regarding the proper concern due to the interests of others"); YAFFE, *supra* note 29, at 117–19. Note that this Article's discussion of the insufficient concern principle here more closely aligns with quality-of-will theorists, who focus on whether a wrongful act manifests an objectionable attitude towards others, in contrast to reasons-responsiveness theorists (which includes Yaffe), who focus on whether a wrongful act manifests

illustrate, consider another iteration of the line-cutter scenario discussed in Part II. Imagine that the line-cutter is aware that his conduct might cause serious injury to both the mother and child, but that he also *really* wants to make sure he makes it into Costco in time to secure toiletries for his family. Under these circumstances, someone who *actually* cares about the well-being of other human beings would view their awareness of a grave risk of harm as a decisive reason to refrain from engaging in the proposed course of conduct, based on the recognition that a mother and child's interests in avoiding life-altering injuries are orders of magnitude greater than the timely acquisition of toiletries. So, when the line-cutter chooses to proceed in the face of this grave risk of harm, it *appears* as though the line-cutter's wrongful behavior is attributable to his failure to care enough about the mother and daughter's well-being.

And yet, to safely (and accurately) make this attribution, there is more we need to know more about the line-cutter's mind. Specifically, when the line-cutter consciously chose to disregard the mother and daughter's well-being, was he fully able to think and act morally? Or was the line-cutter suffering from material impairments in one or more of these psychological domains[171]—for example, serious mental illness,[172] immaturity,[173] and severe emotional distress[174]—which can make it more difficult for people to engage in moral reasoning (i.e., determine right from wrong) or to control their conduct (i.e., conform one's behavior to one's moral judgments).

---

something problematic about how the agent transacts with reasons. YAFFE, *supra* note 29, at 77. It is questionable whether any substantive difference between these two views exists. *Id.* ("To fail to grant the right degree of reason-giving weight to that fact *is* to manifest bad quality of will; and one cannot manifest bad quality of will of the relevant sort unless one also fails to grant appropriate reason-giving weight to the relevant fact about the act, namely that it causes pain."). Nevertheless, I opt for the quality-of-will framing because it is arguably more intuitive than speaking about "mode[s] of transaction with reasons." *Id.*; *see also* MICHAEL MCKENNA, CONVERSATION AND RESPONSIBILITY 57–64 (2012) (discussing and clarifying the meaning of the phrase "quality of will").

171. *See, e.g.*, Ferzan, *supra* note 132, at 2532; *see also* Morse, *supra* note 132, at 62.

172. *See, e.g.*, Westen, *supra* note 29, at 366 ("[A]n insane person who wrongfully kills another because of an unreasonable mistake of fact is excused because, although the insane person commits the *actus reus* of killing an innocent person, he lacks the attitudes of maliciousness, contempt, indifference, disregard, and neglect toward the legitimate interests of others that state-imposed blame represents offenders as possessing.").

173. *See, e.g.*, *id.* at 364 ("Because children are incapable of appreciating those interests in the way adults do, their conduct is incapable of manifesting disparaging attitudes toward those interests—or, at least, incapable of manifesting the kind of malice, contempt, indifference, disregard and neglect that the state expresses when it punishes criminal offenses.").

174. *See, e.g.*, Westen, *supra* note 46, at 156 ("Whether strong emotion consists of the urgent impulse of 'fight or flight' that accompanies anger or the hopeless gloom of depression, strong emotion renders an actor less blameworthy because, while it does not make it impossible, it makes it *much more difficult* for an actor than if he were cool-headed to deliberate about and act upon his settled values."); *see also* Ferzan, *supra* note 135, at 97–98 n.60 ("I also believe that some emotions do prevent us from having full access to our reasons.").

The ability to think and act morally is generally considered to be a precondition for deliberating about and acting upon one's values; therefore, to the extent these rational and volitional capacities are diminished, so too is our basis for attributing wrongful conduct to deficient values.[175] Think of it this way: "If an individual has a mental disability, disturbance, or is overwhelmed by emotion, he cannot reflect meaningfully on his choices."[176] In which case, these impairments may impact the extent to which an action reflects the agent's actual attitudes.[177]

Notice here that in explaining why the mind matters to blame we come across the same four psychological domains discussed in Part II. That is no coincidence; this convergence simply reflects the common philosophical bond that exists between mental states, moral judgment, and human concern. Nevertheless, this convergence is arguably significant: It reveals that the community's sense of psychological blameworthiness is *morally coherent*.[178] That is, a wrongdoer's motivations, awareness of risk, rationality, and volitional control *should* matter to us, given our *ex ante* obligation to care about others, because of what these mental characteristics express: whether a wrongdoer's conduct is attributable to deficient other-regarding values,[179] in contrast to situational factors beyond that person's control (e.g., mental illness, external coercion, and the like).

The minds of others are thus appropriately integral to our blaming practices because of the connection between mental states and social values. And because of that close connection, public expressions of blame necessarily imply the presence of a guilty mind, as a key indicator of socially

175. *See, e.g.*, SCANLON, *supra* note 42, at 153 ("It follows from the way in which blame depends on an agent's reasons that conditions under which an agent acted, such as extreme stress or fear, can affect blame insofar as they affect the degree to which the action reflects the agent's actual attitudes."); Ferzan, *supra* note 135, at 106; Westen, *supra* note 46, at 156.

176. Ferzan, *supra* note 135, at 106.

177. SCANLON, *supra* note 42, at 153. Consider, for example, the import of discovering that the line-cutter was suffering from a deep psychosis at the time of his act that precluded his ability to recognize the moral difference between the mother and child's interest in their bodily integrity versus his own interest in securing toiletries. Or what if one were to learn that the line-cutter was forced to imbibe a narcotic at gunpoint a few moments earlier, which made it impossible for him to conform his behavior to what he might otherwise recognize as the morally right thing to do. In both of these scenarios, the line-cutter's capacity for moral decision-making would be inhibited in a way that precludes us from attributing his conduct to a lack of sufficient concern for the mother and child. In contrast to a deficit of caring, the line-cutter's decision would be attributable to cognitive and volitional hurdles that were beyond his control.

178. Which is to say that it is logical and consistent when viewed in light of our moral commitments.

179. By "other-regarding values," the Article does not mean just *any* moral values—or character traits—that the wrongdoer happens to possess. Rather, the phrase refers to those moral values that speak to the evaluative weight the wrongdoer afforded their own interests in contrast to those of the victim or society that their conduct infringed upon *on that particular occasion*. *See, e.g.*, YAFFE, *supra* note 29, at 25–26; ALEXANDER & FERZAN, *supra* note 29, at 6.

deficient values. Once one understands this basic philosophical insight, the case for limiting blame to those who act with a guilty mind begins to take shape. In the next Section, I focus first on the fairness and societal welfare considerations that support this limitation *in* any public context. Thereafter, I hone-in on the strength of these considerations in the context of the criminal law.

B. Fairness and Societal Welfare Considerations

Why should we refrain from blaming those who lack a guilty mind? Well, for one thing, blaming someone in the absence of a guilty mind would be dishonest, given the social meaning of blame. To denounce someone for failing to sufficiently value societal interests, when in fact they did sufficiently value those interests, expresses "a kind of falsehood."[180] And thus, as Sandy Kadish argues, "to the extent the person is injured by being blamed, [it would be] unjust to him."[181]

The critical point here is less about honestly blaming than it is about blaming fairly. Indeed, justice is arguably the strongest justification for conditioning blame upon the guilty mind.[182] Most often, justice-based arguments are framed through a retributive lens and articulated in terms of *desert*. For example, one could say (and many have said) that in the absence of a guilty mind, one does not deserve to be blamed, in which case doing so would be unjust.[183] However, this also invites an immediate objection: Why *must* someone possess a guilty mind to deserve blame? After all, wrongdoing is also widely considered to be a central ingredient of one's deservingness for blame—so in situations where the harm caused is particularly egregious (e.g., unjustifiably causing someone's death), might this be a reason enough to blame someone?

Academic consensus indicates "no,"[184] however, scholars reach this conclusion in different ways. One common approach is through a

180. Kadish, *supra* note 29, at 264; *see also* T.M. SCANLON, WHAT WE OWE TO EACH OTHER 267 (1998).

181. Kadish, *supra* note 29, at 264.

182. Both here, and throughout the rest of this Article, the phrase "conditioning blame upon the guilty mind" should be understood to mean: (1) only blaming guilty minds; and (2) blaming minds in accordance with their guiltiness.

183. *See, e.g.*, Richard Singer & Douglas Husak, *Of Innocence and Innocents: The Supreme Court and Mens Rea Since Herbert Packer*, 2 BUFF. CRIM. L. REV. 859, 859–62 (1999) ("[C]oncern with whether the conduct of the defendant manifested an evil mind reflects a basic and fundamental principle of justice: Only the blameworthy (guilty), and not the blameless (innocent), [deserve to] be punished."); Garvey, *supra* note 33, at 566 (someone who lacks a guilty mind (because they are ignorant) "deserves pity, not blame . . . the state's censure, [or] the suffering that turns censure into punishment")

184. *See, e.g.*, Stephen J. Morse, *The Moral Methaphysics of Causation and Results*, 88 CALIF. L. REV. 879, 879 (2000) ("[V]irtually all criminal law theorists agree that moral fault is at least a

voluntarism-as-fairness style of argument, which contends that blame can only be deserved in a situation where the person "*could* have done otherwise."[185] Whether this (compatibilist) standard of free will[186] is met is typically understood to rest upon the same kind of psychological capacities (e.g., rationality and volitional control) that ground a guilty mind.[187] Accordingly, a person who engages in wrongdoing but lacks a guilty mind arguably has not exercised the kind of free will from which we can say they were able to do otherwise. Under these circumstances, that person would not deserve to be blamed for their wrongdoing, in which case doing so contrary to desert would be unjust.[188]

It is important to note that retributive arguments for mens rea go beyond the mere injustice of undeserved *blame*. They are equally concerned with the unpleasant consequences that accompany that blame. In the best case, for example, "[t]he sanctions of personal reproach are gradational and relatively mild, ranging from a raised eye-brow, to verbal chastisement, and to social ostracism."[189] However, in a world that is more interconnected than ever, being blamed can also bring with it life-altering repercussions—for example, the loss of friends, social status, and employment, as well as impairments to one's mental health and physical well-being.[190] If, as retributivists contend, the presence of guilty mind is a necessary condition for being subjected to the detriments of blame, then it would be fundamentally unjust to impose this kind of suffering upon someone who lacks a guilty mind.

Conditioning blame upon the guilty mind is a norm about justice, but it may also be a socially beneficial one. For example, some have argued that the norm is a part of a broader network of moral principles and practices that

necessary condition of blame and punishment . . . ."); Singer & Husak, *supra* note 183, at 859–62; Garvey, *supra* note 33, at 566.

185. *See, e.g.*, Kadish, *supra* note 29, at 266 ("It may be said in these cases that the person had no *effective* choice or that no reasonable and upright person *could* have done otherwise."); Vincent Chiao, *Action and Agency in the Criminal Law*, LEGAL THEORY, 2009, at 1.

186. Luis E. Chiesa, *Punishing Without Free Will*, 2011 UTAH L. REV. 1403, 1407.

187. *See, e.g.*, Garvey, *supra* note 33, at 554 ("Even when an actor has committed a criminal wrong, he's not responsible or culpable for his wrongdoing unless, for example, he freely committed the wrong, his wrongdoing was in his control, he could have done otherwise, he acted voluntarily, and so on. An actor doesn't shed his immunity from liability unless what he did was done, we might say more generally, of his own free will."); Brink, *supra* note 41, at 354.

188. *See, e.g.*, Amy J. Sepinwall, *Faultless Guilt: Toward A Relationship-Based Account of Criminal Liability*, 54 AM. CRIM. L. REV. 521, 531 (2017) (summarizing this position as: "[t]o impose blame where we could not have done other than what we did, or for the acts of someone whom we cannot control or have no duty to control, is to hold us responsible for something outside of our agency, and so to treat us more harshly than we deserve." (footnote omitted)).

189. Westen, *supra* note 29, at 327.

190. This is to say nothing of the risk that someone (e.g., the party doing the blaming or a third party) will resort to extra-legal forms of retaliation, such as, for example, the intentional infliction of physical violence.

afford our lives meaning and purpose.[191] The idea is that by respecting the guilty mind, we reaffirm critical ethical values and commitments (e.g., individual autonomy, human dignity, personal identity, and self-worth) that jointly transform a physical universe comprised of atoms into a moral world comprised of persons.[192] Admittedly, these "meaning of life" arguments[193] exist at a relatively high level of abstraction. But for those in search of more tangible societal benefits, there may still be much to recommend conditioning blame upon the guilty mind.

Consider, for example, the prospect that doing so reinforces socially useful messages.[194] One possible message is: sufficiently concern yourself with the interests of others, and you can rightfully expect to avoid being blamed. On this accounting, conditioning blame upon the guilty mind affords those who make reasonable decisions a modicum of security against unwarranted societal condemnation. This is valuable not only because of how destructive blame can be, but also because good intentions are frequently insufficient to prevent one from doing the wrong thing. Strictly limiting blame to those who act with a guilty mind thereby inoculates those who mean well from some of the most aversive consequences when they fall short.

A second possible message communicated by conditioning blame upon the guilty mind is reflected in the converse: Sufficiently concern yourself with the interests of others—or else. A world in which people at least minimally care about one another is not only a nice idea, but a safer and more cohesive place to live. So, to the extent that only blaming those who act with guilty minds makes people less likely to act with them, everyone stands to benefit.

* * *

Thus far, we've been discussing the fairness and general social welfare justifications that support conditioning the *informal* blame we mete out in our

191. *See, e.g.*, Pillsbury, *supra* note 169; Morse, *Inevitable Mens Rea*, *supra* note 132, at 61 ("[T]he requirement of mens rea contributes to the meaning and value of our lives as moral beings.").

192. *See, e.g.*, Daniel Maggen, *Conventions and Convictions: A Valuative Theory of Punishment*, 2020 UTAH L. REV. 235, 265–67; Samuel Scheffler, *The Good of Toleration*, *in* EQUALITY AND TRADITION: QUESTIONS OF VALUE IN MORAL AND POLITICAL THEORY 312, 325 (2010).

193. For some insightful contributions in this area, see, for example, SUSAN WOLF, MEANING IN LIFE AND WHY IT MATTERS (2012); ROBERT NOZICK, *Philosophy and the Meaning of Life*, *in* PHILOSOPHICAL EXPLANATIONS 571 (1981).

194. The idea here is that blame "provides occasion for the wider community to enforce or affirm certain norms and values." Sepinwall, *supra* note 188, at 537; *see, e.g.*, EMILE DURKHEIM, THE DIVISION OF LABOR IN SOCIETY 60–64 (W.D. Halls trans., The Free Press, 1984); David Garland, *Sociological Perspectives on Punishment*, 14 CRIME & JUST. 115, 123 (1991).

social lives on the guilty mind. However, as one turns from the informal to the *formal*, and the social to the *criminal*, the force of the justifications is arguably amplified many times over. This is so due to a key similarity and a few important differences between social and criminal blaming practices.[195]

First, the similarity: The moral judgments rendered by the criminal legal system—namely, conviction and punishment—incontrovertibly constitute a *formal* manifestation of *public* blame.[196] It is widely understood, for example, that a criminal conviction expresses an official judgment of community condemnation, while the sentence attached to it—whether probation, a week of jail, or a lifetime of confinement—denotes the *extent* of that condemnation.[197] The imposition of criminal liability and punishment are thus "public act[s] by which society officially expresses its present moral indignation toward a criminal defendant for violating a criminal prohibition in wrongful disregard of the interests of others."[198] And like the informal blame we mete out in our social lives, the criminal law's formalized judgments of blame carry with them an implicit message of insufficient concern, thereby presupposing the existence of a guilty mind.[199]

195. *See, e.g.*, Westen, *supra* note 29, at 326 ("Legal norms of state-imposed punishment differ from personal norms of interpersonal reproach because, even if the two sets of norms have common origins, the institutions and sanctions of state-imposed punishment differ significantly from those of interpersonal reproach."); Kadish, *supra* note 29, at 265 n.20 ("This, of course, is not to embrace the retributive view that responsibility for law violation itself requires punishment, only that responsibility is necessary, but not sufficient, for punishment."); W.D. ROSS, THE RIGHT AND THE GOOD 60–61 (1930); H.L.A. Hart, *The Presidential Address: Prolegomenon to the Principles of Punishment*, 60 PROC. ARISTOTELIAN SOC'Y 1, 9–10 (1960).

196. Kadish, *supra* note 29, at 289.

197. *See, e.g.*, Westen, *supra* note 29, at 357 ("In reality, the criminal justice system not only officially adjudicates the existence of prohibited conduct but also officially reproaches and condemns the actors whom it finds engaged in it."); JOEL FEINBERG, DOING AND DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY 98 (1970) ("[T]he expression of the community's condemnation is an essential ingredient in legal punishment . . . .").

198. Westen, *supra* note 163, at 188 ("State-imposed condemnation is a moral judgment and, specifically, a moral judgment that the actor disregarded the interests of others. It follows, therefore, that an actor who is condemned for violating a criminal prohibition does not deserve condemnation unless the criminal conduct, in turn, violates interests that are themselves morally legitimate."); Westen, *supra* note 29, at 354.

199. Westen, *supra* note 46, at 151 (asserting that to convict someone of a crime "is to adjudge that, rather than being motivated in his conduct by proper regard for interests that the law seeks to safeguard, the person placed insufficient value on those interests."). Note that, in the criminal law context, condemnatory messages are typically communicated with even greater specificity by virtue of the fact that different offenses exist to safeguard different kinds of interests. As a result, conviction for one offense (e.g. murder) can express something quite different about the quality of an offender's blameworthiness than expressed by a conviction for another offense (e.g., theft). For further discussion of this idea, see *infra* note 245.

While the condemnatory messages expressed through social and criminal blame are similar,[200] the consequences that typically stem from them are not. The stigmatization inherent in a criminal conviction is usually far greater, and reaches a wider audience, than what follows from social blame. We can see this reflected in society's propensity to perceive those convicted of and punished for crimes as less than, or morally inferior to, those who have not.[201] And these are only the informal, social consequences associated with criminal blame. The formal, legal ones—for example, long-term confinement, death, and post-release collateral consequences—are of unfathomable significance to the individual, and exceedingly difficult to correct.[202]

Another important difference between social and criminal blame is that the latter is meted out by the *state*. Arguably, government officials and the varied public institutions they populate are constrained by inviolable "normative limits on the ways in which human beings may be treated," beyond which state action becomes illegitimate.[203] One of the most fundamental limitations is fairness—that is, the state, if it is to threaten formal blame and its attendant sanctions, must afford people a fair opportunity to avoid them.[204] But what does this fair opportunity entail? As with the related

200. A separate but clearly related issue is whether criminal law blaming norms ought to perfectly track our private blaming norms. Westen, *supra* note 29, at 326. The position implicit in this Article is that broad correspondence with private blaming norms is generally a necessary, but not sufficient, condition for justifying a criminal law blaming norm. The differences in consequences between private and criminal law blame offer sound reasons for being far more circumspect about what we criminalize and punish people for than what we might otherwise condemn them for in private. *See id.* at 327 ("[E]ven if lawmakers and the public started with identical senses of wrongdoing and blame, one would expect the official rules of criminal law to differ from the ethical rules of interpersonal relationships.").

201. This lowered social status is perfectly consistent with the social meaning of a criminal conviction and punishment. United States v. Cordoba-Hincapie, 825 F. Supp. 485, 491 (E.D.N.Y. 1993); *see, e.g.*, Richard K. Greenstein, *Toward A Jurisprudence of Social Values*, 8 WASH. U. JURIS. REV. 1, 7 (2015).

202. Westen, *supra* note 29, at 326–27 ("Individuals who reproach one another typically know one another and, if they make mistakes, can correct them; while the institutions of official punishment are state officials and random jurors with no personal knowledge of the events and little ability to correct mistakes.").

203. NICOLA LACEY, STATE PUNISHMENT: POLITICAL PRINCIPLES AND COMMUNITY VALUES 146, 156, 159 (1988); *see also* Brenner M. Fissell, *When Agencies Make Criminal Law*, 10 U.C. IRVINE L. REV. 855, 898 (2020) ("State punishment is a species of coercion, and is thus among the most intrusive forms of state action; even more significantly, though, this coercion takes the form of violence.").

204. *See, e.g.*, HART, *supra* note 29, at 181 ("[U]nless a man has the capacity and a fair opportunity or chance to adjust his behaviour to the law its penalties ought not to be applied to him."); Peter Cane, *Responsibility and Fault: A Relational and Functional Approach to Responsibility*, *in* RELATING TO RESPONSIBILITY: ESSAYS FOR TONY HONORÉ ON HIS EIGHTIETH BIRTHDAY 81, 108 (Peter Cane & John Gardner eds., 2001) ("As agents, we have an interest in freedom of action, in being able to act without incurring the serious penalties and blame that attach to criminal responsibility.").

voluntarism-as-fairness argument discussed earlier, the typical scholarly response focuses on psychological capacities like rationality and volitional control that ground the guilty mind.[205] It therefore follows that when criminal legal actors convict and punish people in the absence of mens rea, they are engaging in an illegitimate exercise of state power.[206]

These differences, between private and criminal blame, significantly strengthen the case for respecting the guilty mind in the legal context. For example, however unfair it may be to blame people who lack mens rea in non-governmental settings, the injustice is that much greater in the criminal legal context, where the condemnation reaches a wider audience, the associated consequences are (typically) more severe, and where those consequences go beyond the state's authority to legitimately impose. Likewise, if conditioning private blame upon the guilty mind is integral to the pursuit of a meaningful life, reinforcing community values, and creating a sense of security from unwarranted condemnation, then these considerations would be particularly pronounced in the criminal legal context, where the judgments are publicly announced, formally endorsed by the state, and delivered on behalf of the "people."[207]

So, in short, while the moral case for respecting the guilty mind is exceedingly strong in informal, private settings, it is stronger yet in the criminal justice arena, insofar as criminal convictions and punishment are concerned. With that in mind, the next and final Part of this Article explores the most important implications of this limitation for U.S. mens rea policy.

---

205. Sepinwall, *supra* note 188, at 530–31; *see, e.g.*, LACEY, *supra* note 203, at 146 ("Both rationality and the capacity for responsible action are thus for liberalism at once factual features of human nature and sources of normative limits on the ways in which human beings may be treated, particularly by political and other public institutions."); HART, *supra* note 29, at 152 (emphasizing "those whom we punish should have had, when they acted, the normal capacities, physical and mental, for doing what the law requires and abstaining from what it forbids, and a fair opportunity to exercise these capacities").

206. *See, e.g.*, Garvey, *supra* note 33, at 546 ("The state gets to decide what makes for a guilty mind when it defines the mental-state elements of a crime, and indeed it should get to decide, as long as it has the authority to do so. But any authority is subject to limits, and the guilty mind as mens rea is one such limit. No state can legitimately punish an actor unless he committed a crime with mens rea.").

207. *See, e.g.*, Henry M. Hart, Jr., *The Aims of the Criminal Law*, 23 L. & CONTEMP. PROBS. 401, 401, 410 (1958). There are, of course, many other utilitarian reasons that could support this limitation on criminal blame. For discussion of the crime control benefits of aligning criminal policy with the community's sense of blameworthiness, understood in terms of promoting perceived legitimacy and voluntary compliance with the law, see generally Paul H. Robinson, *Strict Liability's Criminogenic Effect*, 12 CRIM. L. & PHIL. 411 (2018); ROBINSON, *supra* note 136. For my perspective on this theory of "empirical desert" as it applies to mens rea policy, see Serota, *supra* note 6, at 154–59.

IV. REFORMING THE LAW OF GUILTY MINDS

So far, this Article has focused on the concepts, perspectives, and principles that animate mens rea. In the course of doing so, I've explained what psychological blameworthiness is, which mental states contribute to it, and why criminal policy should respect it. Now it is time to transition from the theoretical to the practical. Assuming one agrees that criminal liability ought to be limited to the psychologically blameworthy, what implications does this have for mens rea policy and reform? That is the focus of this Part.

Section IV.A explains why requiring the government to prove purpose, knowledge, recklessness, or negligence as to every element of an offense is a necessary, but insufficient, form of mens rea protection. The Guilty Minds model reveals diverse ways that contemporary mens rea policies subject insufficiently blameworthy actors to criminal convictions. Because the mismatch between PKRN mens rea and psychological blameworthiness is attributable to the architecture of PKRN mens rea policies, the problem is not one that can be fixed by piecemeal changes to those policies. It is therefore necessary to find a structural solution that can bolster mens rea protections in tandem with the protections afforded by universal culpable mental state requirements.

Section IV.B develops that kind of solution: a generally-applicable "insufficient blameworthiness" affirmative defense built upon the Guilty Minds model. Intended to operate in tandem with PKRN mens rea requirements, an insufficient blameworthiness defense would empower factfinders to dismiss charges in appropriate cases based upon a guided holistic assessment of an accused's mitigating mental states. After providing a statutory codification and detailing its mechanics, I argue that the proposed defense would be accessible to juries and administrable by courts, while finding support in a range of pre-existing doctrines and practices.

A. The Structural Deficiencies of Universal PKRN Requirements

The PKRN approach to mens rea, as originally developed by the Model Penal Code, is widely considered to be the paradigm for protecting the guilty mind through criminal legislation.[208] And that is for good reason: Working to stem the rising tide of strict liability in the mid-twentieth century, the Model Penal Code drafters not only desired to improve the clarity of mens rea doctrine, but also its morality.[209]

208. *See, e.g.*, Serota, *supra* note 6, at 142–43.

209. *See, e.g.*, Andrew Ingram, *Pinkerton Short-Circuits the Model Penal Code*, 64 VILL. L. REV. 71, 72 (2019) ("The belief that criminal liability should not exceed culpability was a basic premise of the drafters of the Model Penal Code.").

Growing concerns with the breakneck speed of economic development and social change during the industrial revolution led to a flood of offenses denying criminal defendants the opportunity to highlight reasonable mistakes and unavoidable accidents—all in order to ease the prosecutor's path to conviction.[210] The Model Penal Code vehemently rejected these strict liability policies, along with the social-control-at-any-cost logic driving them. "Crime does and should mean condemnation," the drafters reasoned, "and no court should have to pass that judgment unless it can declare that the defendant's act was culpable."[211]

The drafters commitment to mens rea is reflected in the Model Penal Code's widely-acclaimed "principle of correspondence," which requires proof of a culpable mental state—whether purpose, knowledge, recklessness, or negligence—for every element of a criminal offense.[212] Implemented through a series of interpretive rules,[213] the drafters believed adherence to universal culpable mental state requirements would effectively protect those whose conduct does not warrant condemnation from being convicted of crimes.[214]

The Model Penal Code's "strict liability abolition" agenda is arguably the most important contribution to mens rea reform efforts of the last century.[215] Through a handful of succinct general provisions, the Model Penal Code provides a legislative framework for requiring proof of culpable mental state across all offenses (and offense elements). Ultimately, for political and ideological reasons I've explored in other work, this substantive mens rea reform blueprint has been far less influential in state legislatures than the conceptual innovations that comprise the PKRN approach to mens rea.[216] But that has not detracted from the blueprint's prominence: Sixty years after completion of the Model Penal Code, the enactment of universal culpable

210. Wechsler, *supra* note 74, at 1439 (observing "the widespread use of strict liability in penal law—not only in the constantly proliferating corpus of the regulatory statutes but even with respect to some of the elements of the more serious offenses, such as bigamy and statutory rape"); David Wolitz, *Herbert Wechsler, Legal Process, and the Jurisprudential Roots of the Model Penal Code*, 51 TULSA L. REV. 633, 670 (2016) ("At the time the Code was being drafted, strict liability crimes were already widespread in state and federal law and were, in fact, increasing along with the growth in regulations more generally.").

211. MODEL PENAL CODE § 2.05 cmt. 1, at 283 (AM. L. INST. 1985).

212. Darryl K. Brown, *Strict Liability in the Shadow of Juries*, 67 SMU L. REV. 525, 527 (2014).

213. *See, e.g.*, MODEL PENAL CODE § 2.02(1), (3), (4) (AM. LAW INST., Proposed Official Draft 1962).

214. MODEL PENAL CODE § 2.05 cmt. 1, at 282 (AM. L. INST. 1985); *see* Wechsler, *supra* note 74, at 1435 ("If fault is to be found with human conduct because it is offensive in its nature, potentialities or consequences, it surely is essential that the actor knew or should have known the facts that give it this offensive character.").

215. *See* Serota, *supra* note 6, at 127–30, 136–37.

216. *See id.* at 130–32.

mental state requirements remains the focal point of mens rea reform discourse, advocacy, and scholarship.[217]

The reconceptualization of the Guilty Minds approach to mens rea developed in this Article indicates that it is time to broaden the conversation. Although requiring the government to prove purpose, knowledge, recklessness, or negligence for every element of an offense is an effective policy mechanism for protecting insufficiently blameworthy actors in many situations, it most certainly does not protect them in all situations. Section IV.A.1 illustrates critical aspects of psychological blameworthiness that contemporary culpable mental state requirements fail to safeguard, even when those requirements are paired with the broadest versions of existing defenses. Thereafter, Section IV.A.2 argues that targeted reforms to pre-existing mens rea doctrines and defenses are incapable of resolving these shortcomings. A structural solution that can work alongside universal culpable mental state requirements is thus necessary.

1. Understanding the Mismatch Between Psychological Blameworthiness and PKRN Mens Rea

To understand what PKRN misses, let us examine what the Guilty Minds model reveals about psychological blameworthiness. First, the blameworthiness of a mind exists at the intersection of four domains of psychological inquiry: motive, risk awareness, rationality, and volitional control. Second, these domains exists on a spectrum in which relevant mental states can strongly mitigate blame, strongly aggravate it, or fall somewhere in between. Because of these characteristics, evaluations of psychological blameworthiness must be *graduated* and *aggregative*.[218] That is, moral judgments about a wrongdoer's mind must be sensitive to the valence—whether positive, negative, or neutral—of each individual psychological domain, and then account for the combined effect of those domains on an actor's overarching blameworthiness.

The PKRN approach to mens rea is structured in a way that conflicts with these realities of moral decision-making. The fundamental problem (further explored below) is that: on the one hand, psychological blameworthiness emerges (or fails to emerge) from a sliding-scale, multi-dimensional analysis of mental states; whereas, on the other hand, PKRN mens rea evaluations are primarily comprised of discontinuous, all-or-nothing rules that address the moral relevance of individual mental states in

217. *Id.*

218. *See* Serota, *supra* note 7, at 1222 (noting the "continuous, graduated judgments of relative blameworthiness expressed in both public opinion surveys and scholarly literature"); Kolber, *supra* note 14, at 857 (observing the "smooth input-output relationship people generally expect between culpability and punishment"); ROBINSON & DARLEY, *supra* note 136, at 208–10.

complete isolation from one another.[219] So long as this structural misalignment exists, insufficiently blameworthy actors are bound to slip through the cracks of mens rea policy.

To illustrate, compare motive's influence on psychological blameworthiness with how the PKRN approach deals with it. As illustrated in Part II, the community's sense of deserved punishment appears to be keenly sensitive to the reasons for which a wrongdoer acts. That is, when someone does a bad thing for a particularly bad reason, their mental state appears to aggravate blame in relative proportion to the moral badness of the reason. Conversely, when someone does a bad thing for a particularly good reason, their mental state appears to mitigate blame in relative proportion to the moral goodness of the reason.

By contrast, motive primarily plays a fixed inculpatory role under the PKRN approach, which bases liability upon whether someone caused a harm prohibited by statute "purposely."[220] That is, for some crimes, it is necessary that the accused's "conscious object" have been to perpetrate the harm expressly prohibited by statute (e.g., causing bodily injury, death, or the destruction of property).[221] Narrowly focusing on this kind of motive,[222] however, can lead to a surface-level analysis that misses the more morally salient reasons for an individual's conduct.[223]

Our earlier discussion of *Crossland v. United States* is instructive.[224] In that case, the court (and relevant law) focused on the shallowest explanation of why Crossland flailed his arms: to resist or injure the police officers harassing him. But there is also a less superficial explanation of why he acted

219. For discussion of an important exception, see *infra* Section IV.B.1 (discussing the gross deviation standard incorporated into the Model Penal Code definitions of recklessness and negligence).

220. MODEL PENAL CODE § 2.02(2)(a) (AM. L. INST., Proposed Official Draft 1962) (defining purpose when applied to the result elements in offense definitions); *see* Hessick, *supra* note 136, at 96–97 (collecting statutes on whether someone engaged in specified conduct for an "unlawful purpose").

221. *See* MODEL PENAL CODE § 2.02(2)(a) (AM. L. INST., Proposed Official Draft 1962).

222. There is wide body of scholarship debating what a motive is, and some debate over whether a purpose actually is a motive, or something different. *See generally* Guyora Binder, *The Rhetoric of Motive and Intent*, 6 BUFF. CRIM. L. REV. 1 (2002) (exploring the historical and philosophical development of motive in criminal law practice and theory). The discussion here follows what I understand to be the common-sense, non-legalistic approach of viewing purposes as a sub-category of motive. *See generally* Douglas Husak, *"Broad" Culpability and the Retributivist Dream*, 9 OHIO ST. J. CRIM. L. 449, 472 (2012) (arguing that the "conceptual question" of "what is motive?" can be avoided "by relying on an intuitive ability to recognize motives"); *see also* Binder, *supra*, at 77–78 (critiquing the inconsistency and incoherence of LaFave's definition of motive—any purpose that does not change an act's character as an offense—which some state courts have adopted).

223. *See, e.g.*, Chiu, *supra* note 136, at 656; Husak, *supra* note 136, at 3; Hessick, *supra* note 136, at 94–95.

224. See *supra* notes 117–119 and accompanying text.

violently toward the police officers: to avoid further unconstitutional, racially-discriminatory abuses. However, this deeper, blame-diminishing motivation was ignored entirely—as is typically the case during the liability stage of criminal proceedings.[225]

I say "typically" because the criminal law occasionally recognizes blame-diminishing motives in the context of general affirmative defenses.[226] For example, the conventional justification of self-defense speaks to circumstances in which causing physical harm in order to prevent physical harm can exculpate, whereas the much rarer lesser-evils defense recognizes the potential exculpatory impact of other socially-legitimate reasons for engaging in wrongdoing.[227] In practice, however, these defenses tend to be extremely restrictive: Among other limitations, they only apply where a person's motives are so good or societally beneficial as to outweigh the harm risked or caused.[228] Where, by contrast, a wrongdoer's motivations are only *partially* mitigating,[229] they typically fall by the wayside in threshold assessments of mens rea—as is illustrated by the *Crossland* case. This creates a large gulf between motive's capacity to diminish psychological blameworthiness and its influence on the liability stage of criminal proceedings.

A comparable gulf exists in how contemporary criminal codes treat rationality and volitional control. As illustrated in Part II, the community tends to view people who find it harder to think or act morally due to impairments for which they are not otherwise responsible (e.g., mental illness or external coercion) as being less blameworthy, in relative proportion to the extent and severity of the impairment. By contrast, the PKRN approach sharply constrains both the impairments that can receive legal recognition and the circumstances in which they are entitled to recognition (namely, when a particular impairment's influence is so strong as to completely negate blameworthiness).

The law of excuses characteristic of the PKRN approach is illustrative. This notoriously narrow area of defense law only recognizes impairments of

225. The reference to the "liability stage of criminal proceedings," both here and throughout this Part, is meant to exclude the sentencing stage, during which the mitigating mental states discussed in this Article (including good motives) may be recognized by courts, whether formally or informally, *see* Hessick & Berman, *supra* note 13, though rarely reliably or predictably, *see* Serota, *supra* note 7.

226. Hessick, *supra* note 136, at 97 ("Motive's fully exculpatory role is perhaps best illustrated by justification defenses.").

227. For discussion of justifications, including self-defense, and a collection of relevant sources, see *supra* note 91.

228. *See, e.g.*, sources cited *supra* note 91.

229. For relevant discussion of partial defenses based on mitigating mental states, see generally Husak, *supra* note 10, and Husak, *supra* note 28.

rationality and volition that stem from a narrow set of sources—namely, mental illness (insanity), the involuntary ingestion of drugs or alcohol (intoxication), and immediate human threats involving grave injury (duress).[230] By contrast, garden variety psychological influences that are equally capable of degrading one's ability to think or act morally—for example, reliving prior trauma,[231] or experiencing tragedy in the present[232]—are generally excluded from the scope of general excuse defenses.

Even within their limited zone of coverage, moreover, general excuse defenses tend to only recognize the most extreme forms of rational and volitional impairment. Consider the *M'Naghten* approach to insanity, which is applied by thirty-four states and the federal government.[233] Under this approach, mental illness is legally relevant when, but only when, it is so severe as to render the person *completely unable* to appreciate the morality of their conduct.[234] Where, by contrast, someone engages in criminal wrongdoing under the influence of severe, rationality-impairing mental illness that falls just a hair short of meeting this high standard, criminal codes effectively deem it immaterial.[235]

A similar dynamic is reflected in the criminal law's other primary excuse defense, duress.[236] Under the majority approach to the subject, a person who, upon being confronted with an immediate and grave threat of violence, *completely* loses their ability to control their conduct is fully exculpated of the crime they were coerced to commit.[237] Where, by contrast, a person is confronted with a hard choice (e.g., assault another or be assaulted yourself) but merely experiences *serious difficulty* in abstaining from intentional wrongdoing, the resulting volitional impairment is—once again—effectively deemed immaterial by our criminal codes.[238] (And this is

230. For discussion of excuse defenses, including mental illness and duress, and a collection of relevant sources, see *supra* notes 89–90.

231. Imagine, for example, that Crossland's child or spouse had been subjected to repeated verbal abuse and harassment in the past by the same offending officers, and that Crossland had lodged a complaint that was ignored without any disciplinary action being taken against the officers.

232. Imagine, for example, that Crossland had just learned about the untimely death of his mother and brother in an unexpected car accident.

233. LAFAVE, *supra* note 89, § 7.2(a).

234. *See id.* § 7.1(a) (explaining that under this test, a mentally ill person will only be excused where they were "laboring under such a defect of reason, from a disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, as not to know he was doing what was *wrong*" (emphasis added)).

235. Kolber, *supra* note 14, at 870.

236. *See, e.g.*, DRESSLER, *supra* note 55, § 23.01; LAFAVE, *supra* note 89, § 9.7.

237. *See, e.g.*, DRESSLER, *supra* note 55, § 23.01; LAFAVE, *supra* note 89, § 9.7.

238. *See, e.g.*, DRESSLER, *supra* note 55, § 23.01; LAFAVE, *supra* note 89, § 9.7.

just one of the many unintuitive thresholds incorporated into contemporary duress doctrine.[239])

Thus, in all of the above instances, as well as many others, the PKRN approach deploys legal tests and doctrines that are chiefly motivated by psychological blameworthiness yet do an exceedingly poor job of aligning criminal liability with it.

2. Solving the Mismatch Between Psychological Blameworthiness and PKRN Mens Rea

In diagnosing the mismatch between psychological blameworthiness and PKRN mens rea, it is important to be clear about the true source of the problem—because it will ultimately bear on fashioning an appropriate solution. Read through a few illustrations of how mens rea doctrine fails to account for the community's sense of blameworthiness, and one might reasonably conclude that both the relevant issues and appropriate responses are fundamentally local—i.e., directly tied to individual mens rea doctrines. If, for example, contemporary formulations of insanity or duress are too narrow or demanding, then those formulations ought to be broadened to more closely mirror the community's sense of psychological blameworthiness.

That may be true,[240] yet it is also true that these kinds of domain-specific reforms only take one so far, given the structural roots of the mismatch. Simply put, the all-or-nothing, domain-specific tests that ground the PKRN approach (and which are integral to its administrative efficiency) will always fail to mirror the community's sense of blameworthiness in at least some circumstances because of the sliding-scale, multi-dimensional nature of moral evaluation.

The mismatch becomes particularly pronounced once one considers that blameworthiness is also a function to the consequences of wrongdoing.[241] All

---

239. Consider, for example, that a duress defense is unavailable, although the accused was wholly unable to control their conduct in response to an explicit threat from the commanding party, if any of the following are true:

(1) The offending threat does not implicate serious bodily injury or death (e.g., the commanding party says, "commit crime *X* or else I will assault you or burn your house down");

(2) the harm threatened is not immediate (e.g., the commanding party says, "commit crime *X* or else I will kill you *tomorrow*); or

(3) the threatened party is commanded to and actually does commit homicide to save themselves.

See, e.g., DRESSLER, *supra* note 55, § 23.01; LAFAVE, *supra* note 89, § 9.7.

240. Indeed, the Model Penal Code's reformulations of these two affirmative defenses are predicated on the idea of expanding coverage to more closely align with the community's sense of psychological blameworthiness. *See supra* notes 89–90.

241. *See, e.g.*, Paul H. Robinson, *The Role of Harm and Evil in Criminal Law: A Study in Legislative Deception?*, 5 J. CONTEMP. LEGAL ISSUES 299, 320 (1994); Paul H. Robinson,

else being equal, for example, we tend to view those who culpably cause, risk, or threaten less serious harms as being less blameworthy than those who cause, risk, or threaten more serious harms as being more blameworthy.[242] Pair this objectivist reality with the comparable subjectivist realities governing the four psychological domains that comprise the Guilty Minds model, and it becomes clear that no amount of tinkering with individual legal doctrine can capture the nuance—or contextuality—of blameworthiness.

Consider a handful of illustrative hypotheticals:

(1) A severely depressed, recently unemployed parent makes a spur of the moment decision to steal a few hundred dollars in groceries from a supermarket in order to feed her hungry children.

(2) A local artist paints a small mural on a privately owned wall in memory of a young woman of color who was recently killed, unjustifiably, by the police on the adjacent sidewalk.

(3) A father who, after being subjected to repeated racial slurs and profanities in the presence of his children for no reason other than the color of his skin, firmly shoves the antagonist to the ground, thereby spraining the antagonist's ankle in the process.

(4) An eighteen-year-old suffering from a moderate developmental disorder assists his older brother's non-violent theft at a convenience store after the older brother threatens to destroy the teenager's computer should he decline to participate in the criminal scheme.

(5) A frequently abused homeless person suffering from moderate mental illness resists arrest after being told he cannot sleep in the park.

(6) A mother who just discovered that her teenage daughter was diagnosed with cancer assaults her daughter's bully after he confronts her with a particularly vicious taunt on the street.

In each of these scenarios, there is no single dimension of the wrongdoers' minds (i.e., motive, risk awareness, rationality, or volitional control) that seems to warrant a categorical exclusion from criminal liability under any PKRN mens rea doctrine. And the nature of the harms threatened, risked, or caused in these scenarios, while far from the most extreme addressed by a criminal code, still seem sufficiently severe to warrant

Prohibited Risks and Culpable Disregard or Inattentiveness: Challenge and Confusion in the Formulation of Risk-Creation Offenses, 4 THEORETICAL INQUIRIES L. 367, 381 (2003); Stephen P. Garvey, *Are Attempts Like Treason?*, 14 NEW CRIM. L. REV. 173, 191 (2011) (stating that from the community's objectivist perspective, "an actor's blameworthiness depends on what he does, and not merely on what he intends to do," in which case "an actor who succeeds in causing harm should be punished more than an otherwise similarly situated actor whose corresponding efforts fail to cause harm").

242. *See, e.g.*, ROBINSON, *supra* note 136, at 178 tbl.17, 180 tbl.18; ROBINSON & DARLEY, *supra* note 136, at ch. 6.

criminalization. And yet, once we aggregate the combined mitigative force of these actor's mental states alongside the comparatively less serious harms they have perpetrated, we may still encounter the absence of blameworthiness sufficient to support a criminal conviction.

That is the core of the mismatch between the PKRN approach to mens rea and psychological blameworthiness, and it is one that contemporary criminal policy, as presently structured, lacks the resources to address. But what has been need not always be. In the final Section of this Article, I propose a structural solution in the form of a statutory insufficient blameworthiness defense.

B. A Legislative Solution: Insufficient Blameworthiness Defense

The Guilty Minds model reveals why even the most steadfast commitment to PKRN mens rea requirements can result in convictions for insufficiently blameworthy actors. More than just a diagnostic tool, however, the Guilty Minds model lays the foundation for a legislative solution guided by three main objectives. First, it should enable the community's sense of blameworthiness to serve as a principled limit on the scope of criminal liability across cases, without regard to subject matter. Second, it should operate without disrupting or requiring significant changes to pre-existing mens rea policies across jurisdictions, including those that hue most closely to the PKRN approach. And third, it should be administrable by courts and juries in a consistent, predictable, and efficient way.

Consistent with these goals, this Article proposes an affirmative "insufficient blameworthiness" defense,[243] which empowers factfinders to

243. This affirmative defense draws inspiration from a few different sources. The first is Judge David Bazelon's proposal that a defendant be excused "if at the time of his unlawful conduct his mental or emotional processes or behavior controls were impaired to such an extent that he cannot justly be held responsible for his act." United States v. Brawner, 471 F.2d 969, 1032 (D.C. Cir. 1972) (Bazelon, C.J., concurring in part and dissenting in part) (emphasis omitted). Developed in 1972 as a better means of dealing with claims of insanity, this defense was intended to "focus the jury's attention on the legal and moral aspects of criminal responsibility, and to make clear why the determination of responsibility is entrusted to the jury . . . ." *Id.*

The second is Alan Michaels' proposal for a "judgmental descriptivism" in the criminal law. Michaels, *supra* note 69, at 58. Writing in 2000, Michaels advocates for an "increased focus on moral blameworthiness" captured through clearly drafted and descriptively precise codifications. *Id.* Pushing back against Herbert Wechsler's recommended codification of homicide (which became the Model Penal Code approach), Michaels argues that an "[a]ffirmation of a commitment to seek precision in defining crimes . . . need not blind us to considerations of blameworthiness." *Id.* at 59 ("Neither should the variety nor the evolutionary character of our moral judgments deter us from making our best efforts to capture them in the criminal law through both legislative and common-law development.").

Third, in an illuminating 2011 paper, Ken Simons raises the possibility of "an 'insufficient culpability' defence, analogous to the lesser evils defence, that the fact-finder would apply on a case-by-case basis." Kenneth W. Simons, *Understanding the Topography of Moral and Criminal*

dismiss charges based upon a structured assessment of an accused's mitigating mental states. The proposed defense, as I envision it, would center around the application of a multi-factor culpability test in which factfinders evaluate the blameworthiness of the accused through the lens of the four domains of psychological inquiry that comprise the Guilty Minds model.

To illustrate, consider the following legislative proposal:

(a) *Insufficient Blameworthiness.* It is an affirmative defense that the defendant's commission of the charged offense is insufficiently blameworthy to warrant the condemnation of a conviction for that offense.

(b) *Relevant Factors.* In determining whether subsection (a) is satisfied, the factfinder shall consider:

(1) whether the person's conduct was motivated by legitimate societal objectives;

(2) whether the person suffered from rational or volitional impairments beyond their control that made it more difficult for them to conform their conduct to the requirements of law;

(3) the substantiality of the harm caused, risked, or threatened by the person's conduct; and

---

Law Norms, *in* PHILOSOPHICAL FOUNDATIONS OF CRIMINAL LAW 248, 250 (R.A. Duff & Stuart P. Green eds., 2011) (noting that "German criminal law incorporates a version of this idea" by requiring the government to prove both mens rea *and* "Schuld (guilt or blameworthiness)"). "On this approach," as Simons explains, "the jury would have the power to reject a criminal prosecution because, considering all of the relevant factors, the defendant is insufficiently culpable to deserve punishment." *Id.* at 251 (cautioning that "[g]ranting this form of extremely broad, untethered discretions to a fact-finder is very powerful medicine").

Fourth, in 2017, an exemplary group of scholars proposed a broad vision of criminal justice reform based on "a common conviction: that the path toward a more just, effective, and reasonable criminal system in the United States is to *democratize* American criminal justice." Joshua Kleinfeld et al., *White Paper of Democratic Criminal Justice*, 111 NW. U. L. REV. 1693, 1693 (2017). Among its many thoughtful and wise recommendations, this White Paper of Democratic Criminal Justice recommends that:

> All crimes carrying a maximum sentence of more than six months should require a showing of moral blameworthiness, where "moral blameworthiness" entails, at a minimum, disregard for the rights or welfare of others or intent to violate the law. The showing of moral blameworthiness may be framed as a component of *mens rea*, a separate element of the offense, an affirmative defense, or in some other fashion, but it should be construed as a question of fact presumptively in the hands of juries, and it should never be established automatically, mechanically, or as a matter of law.

Id. at 1698.

Fifth, in 2019, I developed a *de minimis* provision for the revised D.C. code in my capacity of Chief Counsel for Policy & Planning for the D.C. Criminal Code Reform Commission. *See generally* D.C. CRIM. CODE REFORM COMM'N, FIRST DRAFT OF REPORT #34 – DE MINIMIS DEFENSE (2019), https://ccrc.dc.gov/node/1377776 [hereinafter DRAFT DE MINIMIS DEFENSE]. The insufficient blameworthiness defense proposed here is builds upon and refines that *de minimis* provision. For discussion of what subsequently happened to the revised D.C. code, see *supra* note 109.

(4) the extent to which the person was aware that their conduct would cause, risk, or threaten that harm.

(c) *Burden of Proof*. The defendant has the burden of proof for this affirmative defense and must prove all of its requirements by a preponderance of the evidence.

There is much that could be said about this legislative codification; in what follows, however, I will focus on what I believe to be the four most important (and perhaps least obvious) aspects of the proposed defense.[244]

First, an insufficient blameworthiness defense is intended to complement, but not supplant, conventional mens rea requirements. The proposed provision is built upon the premise that the prosecution should always be required to prove purpose, knowledge, recklessness, or negligence (or some other comparable mental state) beyond a reasonable doubt as to every element of an offense. Legislative enactment of an insufficient blameworthiness defense would not change that. Rather, it would simply establish an additional analytical step in appropriate situations where PKRN mens rea requirements are met. Upon the presentation of legally relevant evidence, the accused would be afforded an opportunity to persuade the factfinder beyond a preponderance of the evidence that commission of the "charged offense"[245] is insufficiently blameworthy to warrant the condemnation of a conviction for that offense, as proscribed in Sections (a) and (c) of the proposed codification.

Second, an insufficient blameworthiness defense is intended to more closely align the outcomes in criminal prosecutions with the lodestar for criminal lawmaking, the community's sense of blameworthiness.[246] Ideally,

244. For additional analysis relevant to understanding the proposed insufficient blameworthiness defense, see the commentary on the *de minimis* provision I drafted for the revised D.C. code. DRAFT DE MINIMIS DEFENSE, *supra* note 243; D.C. CRIM. CODE REFORM COMM'N, FIRST DRAFT OF REPORT #68 – CUMULATIVE UPDATE TO THE REVISED CRIMINAL CODE COMMENTARY SUBTITLE I: GENERAL PART, at 183–90 (2020), https://ccrc.dc.gov/sites/default/files/dc/sites/ccrc/release_content/attachments/First-Draft-of-Report-68-Commentary-Subtitle-I-General-Part.pdf.

245. That this evaluative process is directed toward the "charged offense" recognizes that the quality of blameworthiness sufficient to be convicted of one offense may be different than that required for another. So, for example, a theft conviction expresses that the perpetrator failed to sufficiently value the victim's property interests, whereas a homicide conviction expresses that the perpetrator failed to sufficiently value the victim's interest in their continued human existence. *See, e.g.*, YAFFE, *supra* note 29, at 75; SCANLON, *supra* note 42, at 203. As a result, the quality of blameworthiness sufficient to warrant the condemnation of a conviction for theft is presumably quite different than that of homicide. *See supra* note 199 and accompanying text.

246. This provision is therefore not an invitation for jurors to nullify the pre-existing policy judgments that exist in a criminal code. *See generally* Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 YALE L.J. 677 (1995); Andrew D. Leipold, *Rethinking Jury Nullification*, 82 VA. L. REV. 253 (1996). Indeed, arguably, the proposed defense would limit jury nullification in the long term, by accounting for the community's sense of justice where the legislature expressly invites it. *See* Robinson, *supra* note 71, at 403–04 ("Studies on jury

criminal offenses should, on their face, only extend liability to actors whose conduct warrants the condemnation of a criminal conviction. However, both the limits of language and the complexity of blameworthiness judgments make it all but impossible to draft offenses that meet this admonition without also creating large gaps in coverage.[247] While lawmakers understandably hue toward overinclusion, prosecutors regrettably remain open to exploiting legislative overbreadth wherever they can find it.[248]

To navigate this tension, Section (b) of the proposed defense translates the Guilty Minds culpability model into a multi-factor test, which provides a critical blameworthiness backstop in relevant criminal prosecutions.[249] That backstop asks the factfinder to perform a sliding-scale evaluation in which:

- the more beneficial the societal objectives motivating the defendant's conduct;
- the more difficult the person found it to conform their conduct to the requirements of the law due to rational or volitional impairments that were beyond their control;
- the less substantial the extent of the harm caused, risked, or threatened by the person's conduct; and
- the less aware the defendant was of that harm; then
- the more likely it is that the defendant lacks the quantum of blameworthiness sufficient to warrant the condemnation of a criminal conviction.

Third, this sliding-scale evaluation would be guided, and ultimately constrained, by the community's values. The blameworthiness analysis in Section (b) does not invite the wholly subjective views of individual jurors; instead, it incorporates the same objective, public values-based perspective applied across the criminal law.[250] So, for example, the typical justifiability

---

nullification indicate that jurors frequently exercise their nullification power to circumvent specific rules when they believe that applying them would conflict with broad normative notions of justice. . . . If the criminal code fails to permit moral judgments where appropriate, the system risks being ignored or subverted").

247. *See generally* DRAFT DE MINIMIS DEFENSE, *supra* note 243, at 3 n.1.

248. *See, e.g.*, Ellen S. Podgor, *The Dichotomy Between Overcriminalization and Underregulation*, 70 AM. U. L. REV. 1061, 1065 (2021); Angela J. Davis, *Racial Fairness in the Criminal Justice System: The Role of the Prosecutor*, 39 COLUM. HUM. RTS. L. REV. 202 (2007).

249. Note that Subsection (b)(3) of this multi-factor test incorporates the substantiality of the harm caused, risked, or threatened. As discussed *supra* in Section IV.A.2, this factor plays an important role in community assessments of blameworthiness; therefore, it is incorporated into the proposed defense.

250. *See, e.g.*, LAFAVE, *supra* note 89, § 10.1(d)(4) ("A person with unusual values might think it more important to preserve a valuable painting than to save a human life; but if, faced in an emergency with a choice of saving one of the two, he should choose to destroy the life to save the painting, the court would disagree as to his choice of values and so reject his defense of necessity." (footnote omitted)); United States v. Ashton, 24 F. Cas. 873, 874 (C.C.D. Mass. 1834) ("The law deems the lives of all persons far more valuable than any property . . . ."); *see also* Duncan Kennedy,

analysis deployed in the criminal law does not authorize factfinders to rely on their own (or the defendant's) idiosyncratic values or personal views on particular kinds of activities. Instead, factfinders must perform the requisite balancing of interests—i.e., weighing the social utility of the defendant's conduct against its harmfulness—by "determin[ing] the relative value of the interests at stake from the point of view of the community," while "defer[ring] to judgments of relative value expressed in existing statutes."[251] The blameworthiness analysis required by Section (b) should function under similar constraints, thereby limiting arbitrariness while promoting greater predictability and uniformity.

This is not to say that administration of an insufficient blameworthiness defense would be entirely predictable, uniform, or otherwise free of cost. Tethering the application of general principles to an objective, public values-based perspective does not guarantee that the blameworthiness analysis will be untouched by individual moral idiosyncrasies. And even when the factfinder uniformly embraces the right perspective, some amount of inconsistency is unavoidable in the application of a multi-factor, sliding-scale analysis.[252] Finally, and perhaps most fundamentally, affording criminal defendants a right to compel that analysis in appropriate cases would undoubtedly add to the administrative burden confronting criminal courts.

And yet, if lawmakers are interested in limiting the injustices that inevitably slip through the cracks of hard-and-fast rules,[253] as they should be, these costs are simply the price of doing business with general principles

Form and Substance in Private Law Adjudication, 89 HARV. L. REV. 1685, 1688 (1976) (noting that the application of moral standards require decisionmakers "both to discover the facts of a particular situation and to assess them in terms of the purposes or social values embodied in the standard").

251. Robinson, *supra* note 71, at 450 (emphasis omitted); *see, e.g.*, LaFave, *supra* note 89, § 10.1(d)(4).

252. *See, e.g.*, Russell D. Covey, *Rules, Standards, Sentencing, and the Nature of Law*, 104 CAL. L. REV. 447, 459 (2016) (stating that because standards are "evaluative" and rely on "multi factor or 'totality of the circumstances' tests," they "are thought to be more indeterminate in their application"); Seana Valentine Shiffrin, *Inducing Moral Deliberation: On the Occasional Virtues of Fog*, 123 HARV. L. REV. 1214, 1222 (2010). *See also* Kennedy, *supra* note 250, at 1688 ("[T]he two great social virtues of formally realizable rules, as opposed to standards or principles, are the restraint of official arbitrariness and certainty.").

253. *See, e.g.*, Larry Alexander, *Introduction to the Symposium on the Rationality of Rule-Following*, 42 SAN DIEGO L. REV. 53, 54 (2005) ("Rules achieve their superior ability to provide guidance by being blunt instruments, over and under-inclusive relative to the background moral goals they are meant to achieve."); Robinson, *supra* note 71, at 403 ("[R]ules may further the legality interests of precision and uniformity, but they may distort the proper distribution of criminal sanctions." (footnote omitted)); Covey, *supra* note 252, at 458 ("[R]ules are invariably both over- and under-inclusive; they will apply in some contexts where they should not and will not apply in some contexts where they should.").

sufficiently flexible to address the complexity of blameworthiness.[254] The best one can ask for is a statutory framework that effectively manages the administrative challenges associated with implementing a blameworthiness analysis in a legal forum—and which could therefore be implemented without unduly disrupting or burdening the legal system.[255] Arguably, the proposed legislative codification meets (and indeed, exceeds) these criteria.[256] The balance of this Section is devoted to explaining why.

In what follows, I compare the insufficient blameworthiness defense with two comparable statutory provisions originally recommended by the Model Penal Code and now deployed in numerous U.S. criminal codes: (1) the gross deviation standard for criminal recklessness and negligence; and (2) the *de minimis* defense. I find that blameworthiness analysis proposed here is ultimately clearer, more constrained, and less susceptible to arbitrariness than similar analyses that are regularly conducted by courts and juries across the country. This comparative perspective offers good reason to think that the proposed insufficient blameworthiness defense could be effectively administered.

1. The Gross Deviation Standard for Recklessness and Negligence

Within legal discourse, the Model Penal Code is typically thought of as the paradigm of a descriptive, rules-focused approach to legislation. But the reality is more complicated. There is, as Alan Michaels has observed, an important element of judgmentalism reflected in the Model Penal Code.[257] This includes the Code's reconceptualization of mens rea. Although most aspects of the PKRN approach reflect descriptive, hard-and-fast rules, the Model Penal Code's widely-adopted definitions of recklessness and negligence incorporate a multi-factor blameworthiness analysis comparable to (though ultimately less clear than) that proposed in this Article.[258]

The first clause of the Model Penal Code definitions establish a clear rule of conduct: recklessness prohibits "consciously disregard[ing] a

254. Robinson, *supra* note 71, at 414–15 (stating that although the use of broad moral principles in the criminal law raises legality problems, it is also "essential to the moral force of the criminal law, and to take away the discretion needed for such normative judgments would seriously distort the proper distribution of criminal liability").

255. *See, e.g.*, Michaels, *supra* note 69, at 82–85 (describing a judgmental descriptivist approach to codification); Robinson, *supra* note 71 (laying out a theory of codification that balances considerations of legality, morality, and efficiency); Serota, *supra* note 7, at 1210–13 (discussing codification criteria that effectively promote "legality-related values of democracy, fairness, liberty, and equality").

256. This is the fourth and most practically important aspect of the insufficient blameworthiness defense, following the three prior aspects already discussed.

257. Michaels, *supra* note 69, at 54.

258. MODEL PENAL CODE § 2.02(2)(c)–(d) (AM. LAW INST., Proposed Official Draft 1962).

substantial and unjustifiable risk,"[259] whereas negligence prohibits inadvertently disregarding the same type of risk when the accused "should be aware of" it.[260] However, that clause is then followed by another that asks the factfinder to consider whether: "[t]he risk [was] of such a nature and degree that, considering the nature and purpose of the actor's conduct and the circumstances known to him, its disregard involves a gross deviation from the standard of conduct that a law-abiding person would observe in the actor's situation."[261] This gross deviation standard effectively appends a complex multi-factor blameworthiness analysis onto evaluations for criminal recklessness and negligence—but why? Recognizing that risk-taking is an unavoidable aspect of life, the drafters of the Model Penal Code deemed it wise in prosecutions for recklessness and negligence to require a broader, totality-of-the-circumstances "culpability judgment" from the factfinder to ensure that the accused's conduct actually "justifies condemnation."[262]

While the motivation behind the gross deviation standard is clear, the actual meaning and central mechanics of the blameworthiness analysis it supplies are not. For example, key concepts, including "gross deviation," "law-abiding person" (or, in the event of negligence, "reasonable person"[263]), and "situation" are intentionally left vague by the Model Penal Code, whose drafters trusted the jury to more or less intuit whether the accused deserves to be condemned under circumstances.[264] Less intentional, though just as problematic, is the Code's lack of clarity over how these key concepts are supposed to intersect with one another. The basic mechanics of the resulting blameworthiness analysis have been subject of extensive academic debate for decades.[265]

259. MODEL PENAL CODE § 2.02(2)(c) (AM. LAW INST., Proposed Official Draft 1962) ("A person acts recklessly with respect to a material element of an offense when he consciously disregards a substantial and unjustifiable risk that the material element exists or will result from his conduct.")

260. *Id.* § 2.02(2)(d) ("A person acts negligently with respect to a material element of an offense when he should be aware of a substantial and unjustifiable risk that the material element exists or will result from his conduct.").

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

262. MODEL PENAL CODE § 2.02 cmt. 4 at 241 (AM. LAW INST. 1985).

263. *Id.* § 2.02(2)(d) ("The risk must be of such a nature and degree that the actor's failure to perceive it . . . involves a gross deviation from the *standard of care* that a *reasonable person* would observe in the actor's situation." (emphasis added)).

264. MODEL PENAL CODE § 2.02 cmt. 3, at 237 (AM. LAW INST. 1985) ("There is no way to state this value judgment that does not beg the question in the last analysis; the point is that the jury must evaluate the actor's conduct and determine whether it should be condemned. The Code proposes, therefore, that this difficulty be accepted frankly . . . .").

265. For but a small sampling of relevant legal commentary, see, for example, Stephen P. Garvey, *What's Wrong with Involuntary Manslaughter?*, 85 TEX. L. REV. 333, 341–42 (2006) (providing an overview of the many "difficult interpretive questions"); David M. Treiman, *Recklessness and the Model Penal Code*, 9 AM. J. CRIM. L. 281, 334 (1981); Eric A. Johnson, *Mens Rea for Sexual Abuse: The Case for Defining the Acceptable Risk*, 99 J. CRIM. L. & CRIMINOLOGY

And yet, these notable drafting issues have not put a damper on the popularity of the Model Penal Code's definitions of recklessness and negligence, which have since become the gold standard in U.S. criminal law. For example, "[a]t least 24 state statutes follow" those definitions in reform jurisdictions,[266] while numerous courts in non-reform jurisdictions have relied on them to interpret vague statutes that lack culpable mental state definitions.[267] As a result, the Model Penal Code's definitions of recklessness and negligence have been applied in countless criminal cases, reaching across both time and place. And throughout this extensive legal practice, factfinders appear to have successfully administered the Code's blameworthiness analysis without notable complaint or complication.

This provides reason to believe that the blameworthiness analysis proposed in this Article could be effectively administered. For one thing, that analysis is arguably less vague and more precise than what is injected into prosecutions by the Model Penal Code definitions of recklessness and negligence. For another thing, the Model Penal Code's blameworthiness analysis operates as a required element in any prosecution for recklessness or negligence (i.e., because it is included in the statutory definition for these culpable mental states). Practically speaking, this means that, in appropriate cases, the government possesses the burden to prove *the presence of sufficient blameworthiness beyond a reasonable doubt*. This demanding procedural treatment of blameworthiness is in stark contrast to what is envisioned by the insufficient blameworthiness defense, which only exonerates where the accused can prove *the absence of sufficient blameworthiness beyond a preponderance of the evidence*.

Understood this way, the proposed defense constitutes a clearer, more structured, and less intrusive means of ensuring sufficient blameworthiness in appropriate cases. That the Model Penal Code's vaguer blameworthiness analysis has been successfully applied as an affirmative element in criminal prosecutions without major complaint provides reason to think the defense formulation proposed in this Article could be administered at least as effectively, if not more so.

1, 10–11 (2009); Eric A. Johnson, *Beyond Belief: Rethinking the Role of Belief in the Assessment of Culpability*, 3 OHIO ST. J. CRIM. L. 503, 506 (2006).

266. United States v. Dominguez-Ochoa, 386 F.3d 639, 646 (5th Cir. 2004). State legislatures have, however, made a variety of minor revisions to the Model Penal Code's definitions of recklessness and negligence in the course of enacting them. For discussion of relevant codification trends, see D.C. CRIM. CODE REFORM COMM'N, *supra* note 85, at 65 n.83 (collecting statutory citations and observing that "a majority of reform jurisdictions omit one or more terms and phrases from the gross deviation analysis").

267. *See* LAFAVE, *supra* note 89, pt. 3 § 14.4(a) n.13 ("Even absent such language in the applicable statute, the Model Penal Code formulation is sometimes employed by courts.").

2. Dismissal in the Interests of Justice

A similar conclusion can be drawn from another broad grant of moral discretion that exists in many U.S. criminal codes: statutory provisions authorizing trial courts to dismiss prosecutions in the interests of justice.[268] Legislatures typically grant courts this authority through one of two kinds of statutes. The first, and most prevalent, are statutes which simply authorize trial courts to dismiss criminal prosecutions "in furtherance of justice," with little additional clarification.[269] This barebones approach to codifying an incredibly broad grant of government power essentially affords trial judges unfettered discretion to decide which charges to dismiss and when.[270]

The second form of legislative grant is based on the Model Penal Code's *de minimis* provision, section 2.12. Developed by the Code's drafters to more clearly "articulate criteria for the exercise of [the court's dismissal] power,"[271] this provision authorizes[272] judges to:

> [D]ismiss a prosecution if, having regard to the nature of the conduct charged to constitute an offense and the nature of the attendant circumstances, it finds that the defendant's conduct:
>
> > (1) was within a customary license or tolerance, neither expressly negatived by the person whose interest was infringed nor inconsistent with the purpose of the law defining the offense; or
> >
> > (2) did not actually cause or threaten the harm or evil sought to be prevented by the law defining the offense or did so only to an extent too trivial to warrant the condemnation of conviction; or

268. Roberts, *supra* note 136, at 332.

269. *E.g.*, MINN. STAT. § 631.21 (2021) ("The court may order a criminal action, whether prosecuted upon indictment or complaint, to be dismissed. The court may order dismissal of an action either on its own motion or upon motion of the prosecuting attorney and in furtherance of justice."); *see* Roberts, *supra* note 136, at 332 (collecting statutory citations and finding that "[f]ifteen states and Puerto Rico have enacted statutes that give the courts power to dismiss a prosecution in furtherance of justice"); *see also* Valena E. Beety, *Judicial Dismissal in the Interest of Justice*, 80 MO. L. REV. 629, 631 (2015) (observing judicial dismissal authority).

270. There is one notable exception within the first category of dismissal statutes, New York, which provides judges with a long list of factors to consider. N.Y. CRIM. PROC. LAW § 210.40(a)–(j) (Consol. 2022). The first nine factors explicitly enumerate specific concerns, few of which have anything to do with the actor's culpability or the seriousness of his conduct, whereas the final factor is a general catch-all provision inviting courts to consider "any other relevant fact indicating that a judgment of conviction would serve no useful purpose." *Id.* § 210.40(j).

271. MODEL PENAL CODE § 2.12 cmt. 1 at 402 (AM. L. INST. 1985).

272. Note that although the Model Penal Code states that courts "*shall* dismiss a prosecution," state legislatures have changed this to "may." *See* Watson v. United States, 979 A.2d 1254, 1265 n.15 (D.C. 2009) (emphasis added).

> (3) presents such other extenuations that it cannot reasonably be regarded as envisaged by the legislature in forbidding the offense.[273]

This *de minimis* provision, currently applied in a handful of jurisdictions,[274] was animated by concerns and objectives quite similar to those behind the insufficient blameworthiness defense. For example, Model Penal Code section 2.12 was intended to provide a critical judicial backstop when overzealous criminal prosecutors exploit the inevitable overbreadth that exists in general criminal legislation.[275] Given the difficulty of drafting criminal offenses that solely apply to those who warrant the condemnation of conviction, the drafters of the Model Penal Code deemed it wise to empower courts to ensure that criminal convictions "reflect the proper level of the defendant's culpability."[276] In affording courts the power to make these discretionary culpability judgments, however, the drafters of the Model Penal Code also believed more guidance and constraint was necessary than what the barebones dismissal statutes previously discussed offer.[277] Section 2.12 thus seeks to clearly codify, in advance, the situations in which the judicial dismissal power "reasonably can be used."[278]

While the goals animating Model Penal Code section 2.12 are laudable, the provision arguably fails to accomplish them—in ways that ultimately reveal the virtues of the insufficient blameworthiness defense. For one thing, the Code's *de minimis* provision, although motivated by a desire to protect

273. MODEL PENAL CODE § 2.12 (AM. L. INST., Proposed Official Draft 1962). For a sense of the range of conduct to which *de minimis* statutes apply, consider the following dismissals collected in Stanislaw Pomorski,*On Multiculturalism, Concepts of Crime, and the "De Minimis" Defense*, 1997 B.Y.U. L. REV. 51, 90 n.137:

> New Jersey v. Bazin, 912 F. Supp. 106 (D.N.J. 1995) (verbal harassment); State v. Zarrilli, 523 A.2d 284 (N.J. Super. Ct. Law Div. 1987) (taking a single sip of beer by an underage boy attending a church function); State v. Smith, 480 A.2d 236 (N.J. Super. Ct. Law Div. 1984) (shoplifting three pieces of bubble gum worth 15 [cents]); State v. Nevens, 485 A.2d 345 (N.J. Super. Ct. Law Div. 1984) (taking fruit from the premises of a buffet-type restaurant after paying for the meal); Commonwealth v. Moll, 543 A.2d 1221 (Pa. Super. Ct. 1988) (damaging a drainage pipe belonging to the town to prevent flooding of the defendant's land (mischief)); Commonwealth v. Jackson, 510 A.2d 1389 (Pa. Super. Ct. 1986) (riot and failure to disperse by prison inmates upon official order); Commonwealth v. Houck, 335 A.2d 389 (Pa. Super. Ct. 1975) (verbal harassment—calling the victim on the phone "morally rotten" and "lower than dirt").

274. Four states and one territory have adopted this provision. *See* N.J. STAT. ANN. § 2C:2-11 (West 2015); HAW. REV. STAT. § 702-236(1) (2016); ME. REV. STAT. tit. 17-A, § 12 (2006); 18 PA. STAT. AND CONS. STAT. ANN. § 312 (West 2006); 9 GUAM CODE ANN. § 7.67 (2014); *see also* Roberts, *supra* note 136, at 336 ("No state has both a *de minimis* and an in furtherance of justice statute: rather, these are alternative ways of serving a similar function.").

275. MODEL PENAL CODE § 2.12 cmt. 1 at 402 (AM. L. INST. 1985); *see supra* note 247 and accompanying text (discussing the problem).

276. MODEL PENAL CODE § 2.12 cmt. 1 at 402 (AM. L. INST. 1985).

277. *See id.*

278. *Id.*

non-culpable actors in a predictable way, says nothing about how considerations of mens rea are supposed to factor into the *de minimis* analysis.[279] Absent legislative clarification, judges have had to resolve the issue for themselves.[280] Generally speaking, courts in relevant jurisdictions seem to embrace the full spectrum of psychological blameworthiness as being potentially relevant to the *de minimis* analysis.[281] But judicial evaluation of "the what, why and how of defendant's intent" in *de minimis* litigation also appears to occur free-form, in the absence of explicit legal guidance on how to account for the various dimensions of psychological blameworthiness.[282] The insufficient blameworthiness defense proposed in this Article, by contrast, provides that guidance in the form of a multi-factor test that clearly addresses which mental states matter (and how).

Equally important is who receives that guidance—and thus is ultimately tasked with conducting the blameworthiness analysis. The proposed defense effectively ask juries to determine which individuals warrant the condemnation of a criminal conviction. By contrast, both Model Penal Code section 2.12 and the barebones dismissal statutes effectively ask judges to conduct this blameworthiness analysis. But that's arguably the wrong institution to ask—for at least two reasons.

First, blameworthiness evaluations call for the moral conscience of the community. Juries, as "the community's sense of justice," are arguably the institution best situated to provide that conscience,[283] which is why juries

279. That is true of each of the three subsections that comprise Model Penal Code section 2.12. *See* Pomorski, *supra* note 273, at 94–98.

280. *Id.* at 97.

281. *Id.* (collecting cases); *see also* Roberts, *supra* note 136, at 375–76 ("Again and again, one finds judges moved to dismiss in light of their assessment of defendants' motives. When those motives are ones esteemed as noble—when, for example, they are focused on the welfare of children—courts show no hesitation in deeming motive a ground for dismissal." (footnotes omitted)).

282. State v. Cabana, 716 A.2d 576, 579 (N.J. Super. Ct. Law Div. 1997) (emphasis added). Interpreting its state *de minimis* provision, New Jersey courts have developed a judicially created framework under which judges consider among other factors:

> (a) Defendant's background, experience and character as indications of whether he or she knew or should have known the law was being violated;
>
> (b) Defendant's knowledge of the consequences of the act . . . and
>
> (h) Any other information which may reveal the nature and degree of culpability.

State v. Halloran, 141 A.3d 1216, 1220 (N.J. Super. Ct. Law Div. 2014) (citing State v. Zarrilli, 523 A.2d 284, 286 (1987)); *see also Cabana*, 716 A.2d at 578 (finding that New Jersey's *de minimis* statute clearly "contemplates" a "threshold consideration of criminal culpability" which is "dependent upon the state of mind of the actor and [requires] a fact-sensitive analysis on a case by case basis"). Hawaii courts apply a similar approach. *See* State v. Park, 525 P.2d 586, 591 (Haw. 1974).

283. *See, e.g.*, Robinson, *supra* note 71, at 460; David L. Bazelon, *The Morality of the Criminal Law*, 49 S. CAL. L. REV. 385, 396 (1976) (noting that, as the community is both the beneficiary and the victim of the insanity defense, the community, "through its traditional surrogate, the jury,"

(rather than judges) are most often relied upon to administer culpability-focused doctrines and defenses (e.g., insanity and duress).[284] Judges, by contrast, are arguably one of the persons least able to reliably reflect the community's moral conscience in legal decision-making given their traditional background, training, and socialization (among other reasons).[285]

Second, blameworthiness evaluations require an institutional actor willing to serve as a counterweight to overly aggressive prosecutors.[286] Juries may, once again, may a better fit for the role. Throughout history juries have provided an effective (though imperfect) "counterforce against overly harsh laws or overzealous prosecutors whose actions transgress the people's sense of justice."[287] For example, whether pushing back against prosecutorial overreach by applying legally-recognized excuses, or through extra-legal acts of nullification, juries have proven themselves willing to police the boundaries of blameworthiness in criminal cases.[288] The judicial propensity to stand in the way of prosecutorial overreach, by contrast, is arguably more dubious. After all, courts have played a critical role in the birth of mass incarceration in the United States, and they continue to take actions which help to sustain in in the present.[289]

---

should "assess criminal responsibility in light of community standards"). Note that recent psychological research suggests "people are relatively calibrated and even-handed in utilizing evidence that either amplifies or mitigates blame." Andrew E. Monroe & Bertram F. Malle, *People Systematically Update Moral Judgments of Blame*, 116 J. PERSONALITY & SOC. PSYCH. 215, 215 (2019). The idea is that "[t]he requirement for warrant and the potential social cost of blaming may motivate people to be relatively careful in attending to available blame-relevant information, including agents' intentionality and mental states, their causal contributions to an outcome, and even counterfactuals about the preventability of the outcome." *Id.*

284. H.R. REP. NO. 1076-90, at 8 (1968) ("[T]he jury is designed not only to understand the case, but also to reflect the community's sense of justice in deciding it.").

285. Robinson, *supra* note 71, at 417.

286. Watson v. United States, 979 A.2d 1254, 1265–66 (D.C. 2009).

287. Ethan J. Leib, Michael Serota & David L. Ponet, *Fiduciary Principles and the Jury*, 55 WM. & MARY L. REV. 1109, 1137 (2014) ("Historically, this has meant that juries sometimes ignored the Fugitive Slave Act or refused to issue guilty verdicts during Prohibition, just as today they may selectively apply drug laws they perceive as unfair.").

288. *Id.; see, e.g.*, Stephanos Bibas, *Transparency and Participation in Criminal Procedure*, 81 N.Y.U. L. REV. 911, 928 (2006) (noting that judges can be more punitive than the general public); JULIAN V. ROBERTS & LORETTA J. STALANS, PUBLIC OPINION, CRIME, AND CRIMINAL JUSTICE 210–12 (1997) (observing that lay decisionmakers can impose sentences that are more lenient than those imposed by judges).

289. *See, e.g.*, Jed S. Rakoff, U.S. Dist. Judge, S. Dist. N.Y., Speech at Harvard Law School Conference: Mass Incarceration and the "Fourth Principle" (Apr. 10, 2015), *in* Gabe Friedman, *Judge Rakoff Speaks Out at Harvard Conference: Full Speech*, BLOOMBERG (Apr. 13, 2015, 1:04 PM), https://bol.bna.com/judge-rakoff-speaks-out-at-harvard-conference-full-speech/ ("[F]or too long, too many judges (including me) have been too quiet about an evil of which we are ourselves a part: the mass incarceration of people in the United States today."); Mark Osler & Judge Mark W. Bennett, *A "Holocaust in Slow Motion?" America's Mass Incarceration and the Role of Discretion*, 7 DEPAUL J. FOR SOC. JUST. 117, 120 (2014) (observing the role of trial court judges in contributing to mass incarceration at the federal level).

That is why, although the goals animating dismissal statutes are admirable, they fall short of achieving them. These statutes fail to supply clear guidance on how to limit liability to sufficiently blameworthy actors. And they look to the wrong decisionmaker to enforce those limits. These shortcomings do not mean that we should withhold from judges the authority to dismiss charges, and there certainly are good reasons to enact (or preserve) dismissal statutes, flaws notwithstanding.[290] However, by reflecting on those flaws, we come to see the insufficient blameworthiness defense's comparative virtues and ultimate indispensability for protecting individuals that do not warrant the condemnation of a criminal conviction.

* * *

There is thus a powerful administrative case that supports the legislative enactment of an insufficient blameworthiness defense. And that case matters—not only because administrability is the key to ensuring that a criminal policy does what it's supposed to do, but also because a criminal policy that cannot be effectively administered is unlikely to actually be enacted. In a world where lawmakers often seek policy guidance from those involved in the daily practice of criminal law (most often, prosecutors, though also courts and defense attorneys, too), the success of a legislative proposal is only as great as its proponents' ability to successfully address concerns about implementation. The final Section of this Article has attempted to address those concerns by illuminating the mechanics of an insufficient blameworthiness defense, the precedents that exist for employing it, and the defense's relative strengths as a mechanism for bolstering mens rea protections in U.S. criminal codes.

CONCLUSION

On the whole, the transition from the common law's vision of mens rea to the Model Penal Code's was one of progress. Replacing the vague and moralistic conception of culpability that is characteristic of the Guilty Minds approach with PKRN's more precise and legalistic understanding has yielded important benefits in jurisdictions across the country. But there is more to the story of mens rea's evolution than is often appreciated. By divorcing mens rea from its moral foundations, the PKRN approach has also exacted underappreciated costs on U.S. mens rea policy. In this Article, I've attempted to spotlight those costs and demonstrate how a more refined understanding of the Guilty Minds approach can be used to address them.

290. *See generally* Pomorski, *supra* note 273.

Synthesizing experimental research, the Article transformed the common law understanding of mens rea into a four-part model of psychological blameworthiness rooted in the community's sense of justice. Building on contemporary criminal law theory, the Article then explained why this reconceptualization of mens rea ought to constrain the breadth of criminal liability in U.S. codes. The Guilty Minds model revealed structural flaws in contemporary mens rea policies that violate this constraint. And that model also provided the foundation for a statutory solution: an insufficient blameworthiness defense, which would empower factfinders to dismiss charges based upon a structured assessment of an accused's mitigating mental states.

Could this defense be successfully administered? There is good reason to think so. Although the proposed codification is novel, asking criminal justice actors to holistically analyze blameworthiness is not. And the blameworthiness analysis proposed in this Article is arguably clearer and more accessible than the comparable analyses that courts and juries around the country successfully perform. All told, an insufficient blameworthiness defense offers all jurisdictions—including those that hue most closely to the Model Penal Code's recommendations—with a promising new pathway for criminal law reform.[291]

291. This Article focused on the moral philosophical case for bolstering mens rea protections in U.S. criminal codes, and the administrative case for using an insufficient blameworthiness defense to do so. However, this is not intended to exhaust relevant justifications. For discussion of the public safety, decarceration, racial justice, and political considerations that support undertaking mens rea reform in a time of mass incarceration, see Serota, *supra* note 6.