

# Oklahoma Law Review

---

Volume 68 | Number 3

---

2016

## Belief States in Criminal Law

James A. Macleod

Gibson, Dunn & Crutcher LLP, [jammacleod@gmail.com](mailto:jammacleod@gmail.com)

Follow this and additional works at: <https://digitalcommons.law.ou.edu/olr>



Part of the [Criminal Law Commons](#)

---

### Recommended Citation

James A. Macleod, *Belief States in Criminal Law*, 68 OKLA. L. REV. 497 (2016),  
<https://digitalcommons.law.ou.edu/olr/vol68/iss3/2>

This Article is brought to you for free and open access by University of Oklahoma College of Law Digital Commons. It has been accepted for inclusion in Oklahoma Law Review by an authorized editor of University of Oklahoma College of Law Digital Commons. For more information, please contact [darinfox@ou.edu](mailto:darinfox@ou.edu).

# BELIEF STATES IN CRIMINAL LAW

JAMES A. MACLEOD\*

*Belief-state ascription—determining what someone “knew,” “believed,” was “aware of,” etc.—is central to many areas of law. In criminal law, the distinction between knowledge and recklessness, and the use of broad jury instructions concerning other belief states, presupposes a common and stable understanding of what those belief-state terms mean. But a wealth of empirical work at the intersection of philosophy and psychology—falling under the banner of “Experimental Epistemology”—reveals how laypeople’s understandings of mens rea concepts differ systematically from what scholars, courts, and perhaps legislators, have assumed.*

*As implemented, mens rea concepts are much more context-dependent and normatively evaluative than the conventional wisdom suggests, even assuming that jurors are following jury instructions to the letter. As a result, there is less difference between knowledge and recklessness than is typically assumed; jurors consistently “over”-ascribe knowledge to criminal defendants; and concepts like “belief,” “awareness,” and “conscious disregard” mean different things in different contexts, resulting in mens rea findings systematically responsive to aspects of the case traditionally considered irrelevant to the meaning of those terms.*

*This Article provides the first systematic account of the factors driving jurors’ ascriptions of the specific belief states criminal law invokes. After surveying mens rea jury instructions, introducing the Experimental Epistemology literature to the legal literature on mens rea, and examining the implications of that literature for criminal law, this Article considers ways to begin bridging the surprisingly large gap between mens rea theory and practice.*

## Table of Contents

I. Introduction .....	498
II. Belief States in Criminal Law .....	504

---

\* Associate, Gibson, Dunn & Crutcher LLP. For helpful comments and suggestions, I thank James Beebe, George Fletcher, Stephen Garvey, Eric Johnson, Brian Leiter, Richard McAdams, Josh Parker, Samuel Pillsbury, Marci Rozen, Alex Sarch, Joseph Shin, Ken Simons, Lior Strahilevitz, David Treiman, John Turri, Sofia Vickery, Ingram Weber, Maggie Wittlin, and participants in the Annual Experimental Philosophy Conference at the University of Buffalo and the Legal Scholarship Workshop at the University of Chicago. I also thank my parents and Norah.

A. The Model Penal Code .....	504
1. Knowingly .....	505
2. Recklessly and Negligently .....	508
B. The Federal Criminal Code .....	511
III. An Empirically Informed Account of Belief State Ascription .....	514
A. Introducing Experimental Epistemology .....	514
1. The Side-Effect Effect .....	515
2. The Epistemic Side-Effect Effect .....	520
B. Bringing the Data Together: A Descriptive Account of Belief State Ascription .....	523
1. Ascribing Belief .....	523
a) The Norm-Violation / Belief-Ascription Heuristic .....	523
b) Different Kinds of Belief: Occurrent and Dispositional, Thick and Thin .....	524
2. Ascribing Knowledge .....	529
a) Pragmatic Features of the Agent's Situation .....	530
b) Normative Valence of the Course of Action at Issue .....	533
IV. Surprising Divergences Between Mens Rea Theory and Practice: What They Are, Why They Matter, and How to Begin Addressing Them .....	534
A. The Vanishing Distinction Between Knowledge and Recklessness .....	535
B. A Peculiar Difference Between Willful Ignorance and Knowledge .....	542
C. Implicit, and Outcome-Determinative, Differences in "Belief" and Its Relatives .....	546
D. The Vanishing Distinction Between Description and Judgment .....	548
E. The Need for Greater Scrutiny of Mens Rea Jury Instructions ...	550
V. Conclusion .....	553

### I. Introduction

The law often requires fact-finders to use circumstantial evidence to determine another's mental state. These mental states fall into two basic categories: desire states (such as intent, purpose, and indifference) and belief states (such as knowledge, belief, and ignorance).<sup>1</sup> Legal scholarship

---

1. See Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV. 463, 476 (1992) [hereinafter Simons, *Rethinking*]. The "desire states" label is a bit of a misnomer: one can *intend* to do something without *desiring* to do it, for example, as happens when one is forced to choose between courses of action one dislikes. Still, the label is prevalent in the legal literature and suffices for present purposes.

has examined how laypeople ascribe *desire* states,<sup>2</sup> but it has largely ignored how fact-finders ascribe *belief* states. This gap in the legal literature is surprising given the central role belief-state ascription plays in fact-finding throughout the law.

Consider substantive criminal law. A defendant's guilt or innocence, along with his degree of punishment, frequently turns on his belief state—whether he was “aware” of something, “believed” something, “consciously disregarded” something, “knew” something, etc. The definitions of all four of the Model Penal Code's (MPC) mens rea categories—purpose, knowledge, recklessness, and negligence—explicitly invoke the defendant's belief state, as do hundreds of provisions throughout federal and state criminal codes.<sup>3</sup> And a defendant's punishment differs significantly depending on whether, for example, the defendant was “aware” of a risk (and hence was reckless, rather than negligent),<sup>4</sup> or whether the defendant

---

2. E.g., Pam A. Mueller et al., *When Does Knowledge Become Intent? Perceiving the Minds of Wrongdoers*, 9 J. EMPIRICAL LEGAL STUD. 859 (2012); Janice Nadler & Mary-Hunter McDonnell, *Moral Character, Motive, and the Psychology of Blame*, 97 CORNELL L. REV. 255 (2012); Thomas Nadelhoffer, *Criminal Law, Philosophy, and Psychology: Working at the Crossroads*, in OXFORD STUDIES IN PHILOSOPHY OF LAW, VOLUME 1 276-82 (Leslie Green, Brian Leiter eds. 2011) [hereinafter Nadelhoffer, *Criminal Law*]; Julia Kobick, *Discriminatory Intent Reconsidered: Folk Concepts of Intentionality and Equal Protection Jurisprudence*, 45 HARV. C.R.-C.L. L. REV. 517 (2010); Julia Kobick & Joshua Knobe, *How Research on Folk Judgments of Intentionality Can Inform Statutory Analysis*, 75 BROOK. L. REV. 409 (2009); Lawrence M. Solan, *Blame, Praise, and the Structure of Legal Rules*, 75 BROOK. L. REV. 517 (2009); Thomas Nadelhoffer, *Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Juror Impartiality*, 9 PHILOSOPHICAL EXPLORATIONS 203 (2006) [hereinafter Nadelhoffer, *Bad Acts*]; Bertram F. Malle & Sarah E. Nelson, *Judging Mens Rea: The Tension Between Folk Concepts and Legal Concepts of Intentionality*, 21 BEHAVIORAL SCI. & L. 563 (2003).

3. E.g., sources cited *infra* notes 12, 24, 45-56. Moreover, some recent legislative proposals, as well as criminal law scholarship, advocate a mens rea schema with yet *greater* emphasis on belief states such as awareness of wrongdoing. E.g., H.R. 4002, proposed Nov. 16, 2015 (proposed “Criminal Code Improvement Act of 2015” would add to federal criminal provisions a requirement that, “if the offense consists of conduct that a reasonable person in the same or similar circumstances would not know, or would not have reason to believe, was unlawful, the Government must prove that the defendant knew, or had reason to believe, the conduct was unlawful.”); Gideon Yaffe, *A Republican Crime Proposal That Democrats Should Back*, N.Y. TIMES, Feb. 12, 2016 (supporting passage of the Criminal Code Improvement Act of 2015); Samuel W. Buell, *Culpability and Modern Crime*, 103 GEO. L.J. 547 (2015); see also Douglas Husak, “Broad” Culpability and the Retributivist Dream, 9 OHIO ST. J. CRIM. L. 449, 480-84 (2012).

4. See MODEL PENAL CODE § 2.02(2) (AM. LAW INST. 1962) [hereinafter MPC]; Douglas Husak, *Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting*, 5 CRIM. L. & PHIL. 199, 207 (2011) (“It is scandalous that the scholarly debate

was not merely “aware” of a risk but actually “knew” that a bad outcome would result from his action.<sup>5</sup>

Despite the central role these and other belief-state terms play in law,<sup>6</sup> a fundamental question remains unanswered: what do they mean? More specifically, when the criminal law asks jurors to determine whether a defendant “knew” something, “believed” something, “consciously disregarded” something, etc., and when jurors faithfully implement the instructions they are given, what functional understanding of those belief-state terms are they employing?

This Article is the first sustained attempt at answering that question. Previous accounts of modern mens rea concepts have too often glossed over it, relying on supposedly commonly shared intuitions about the meaning of the terms at issue,<sup>7</sup> philosophical analyses of the relevant terms,<sup>8</sup> or

---

about the justifiability of penal liability for negligence has not made greater efforts to understand the nature of awareness.”).

5. See *infra* notes 168-71 and accompanying text.

6. Belief states are central to many more issues, in criminal law and elsewhere, than are apparent at first glance. For example, sticking with criminal law, even where the pertinent statute calls for *desire*-state ascription and does not explicitly call for belief-state ascription, courts sometimes explicitly instruct that a belief state satisfies the desire-state requirement—e.g., if the jury finds that the defendant possessed *knowledge* that his action would cause a given outcome, this may suffice for a finding that the defendant *intentionally* caused that outcome. See, e.g., *Rosemond v. United States*, 134 S. Ct. 1240, 1248-49 (2014). And even where courts do not so instruct, recent empirical research indicates that a jury’s intentionality ascription will likely track its knowledge ascription. In other words, with or without an instruction from the court, the jury will typically deem knowledge sufficient for intent. See Mueller, *supra* note 2, at 860; *infra* Section III.A.1. Nor does belief-state ascription merely influence ascription of other mental states. The conclusions that fact-finders make about a person’s belief state sometimes also impact fact-finders’ determination of an act or omission’s degree of *causal influence*—i.e., degree to which the act or omission caused a particular outcome—an attribute traditionally thought not to implicate mental state ascription. See Christopher Hitchcock & Joshua Knobe, *Cause and Norm*, 106 J. PHIL. 587, 602-05 (2009); Nadler & McDonnell, *supra* note 2, at 287; Briar Helen Moir, *Judgments in Causal Chains: The Impact of Positive and Negative Motives and Outcomes on Lay Attributions* (2014) (unpublished Ph.D. dissertation, Victoria University of Wellington), <http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/3277/thesis.pdf?sequence=2> (reviewing the literature). In sum, a growing body of legal doctrine and empirical literature places belief states at the center of many more legal issues than had previously seemed to be the case.

7. See e.g., Eric A. Johnson, *Knowledge, Risk, and Wrongdoing: The Model Penal Code’s Forgotten Answer to the Riddle of Objective Probability*, 59 BUFF. L. REV. 507, 513 (2011) (“There is a difference, of course, between what the actor knew and what she believed; in order to qualify as knowledge, a belief must be (at the very least) true and justified.”); Anthony M. Dillof, *Transferred Intent: An Inquiry into the Nature of Criminal*

appellate case law, statutory structure, official code commentaries, and other sources to which jurors are not privy.<sup>9</sup> In contrast, this Article's method is to examine actual mens rea jury instructions, and then, to the extent they leave relevant issues ambiguous, to draw on empirical research concerning how laypeople are likely to resolve those ambiguities.<sup>10</sup>

---

*Culpability*, 1 BUFF. CRIM. L. REV. 501, 529 (1998) (arguing that an expert shooter *A*, who aims for *B* and is justified in believing he has shot and killed *B*, but who, it turns out, actually shot and killed *C*, did not “know he has killed someone” because “[i]ntuitively, we would say” that he does not know that he has killed someone); Kenneth W. Simons, *Should the Model Penal Code's Mens Rea Provisions Be Amended?*, 1 OHIO ST. J. CRIM. L. 179, 187 (2003) [hereinafter Simons, *Should the Model*] (“[I]f ‘high probability’ is greater than 50%, then the definition [of willful blindness, found in MPC § 2.02(7)] as a whole appears to be incoherent,” because you cannot “actually believe that a fact does not exist, when you know that more likely than not, it does exist”); David Luban, *Contrived Ignorance*, 87 GEO. L.J. 957, 961 (1999) (“[K]nowledge does require belief—I can hardly be said to know something if I don't even believe it”); Christopher Slobogin, *A Rational Approach to Responsibility*, 83 MICH. L. REV. 820, 836 (1985) (reviewing MICHAEL S. MOORE, LAW AND PSYCHIATRY: RETHINKING THE RELATIONSHIP (1989)) (distinguishing between knowledge of the consequences of one's actions and “practical[] certain[ty]” regarding those consequences); *United States v. Jewell*, 532 F.2d 697, 700 (9th Cir. 1976) (en banc) (“[I]n common understanding one ‘knows’ facts of which he is less than absolutely certain. To act ‘knowingly,’ therefore, is not necessarily to act only with positive knowledge . . . .”); *United States v. Golomb*, 811 F.2d 787, 792 (2d Cir. 1987) (“Knowledge and belief are very different mental states; knowledge implies a much higher degree of certainty.”); *Cheek v. United States*, 498 U.S. 192, 207-08 (1991) (Scalia, J., concurring) (“It is quite impossible to say that a statute which one believes unconstitutional represents a ‘known legal duty.’”).

8. See, e.g., Douglas N. Husak & Craig A. Callender, *Wilful Ignorance, Knowledge, and the “Equal Culpability” Thesis: A Study of the Deeper Significance of the Principle of Legality*, 1994 WIS. L. REV. 29, 47 n.72, 51 (using “conceptual analysis,” “retain[ing] the common opinion that” knowledge requires true belief that is justified, and concluding that “[s]ome kind of justificatory condition is necessary for both the legal and philosophical senses of knowledge”); Ira P. Robbins, *The Ostrich Instruction: Deliberate Ignorance as a Criminal Mens Rea*, 81 J. CRIM. L. & CRIMINOLOGY 191, 220-27 (1990) (“Like the philosophical notion of knowledge, criminal knowledge requires certainty and a corresponding absence of doubt. . . . Therefore, one ‘knows’ something only if he or she is certain of it.”).

9. See Samuel H. Pillsbury, *Crimes of Indifference*, 49 RUTGERS L. REV. 105, 207-08 (1996) (explaining that jury instructions have not “received the scholarly attention they deserve,” since “[s]cholarly analysis of mens rea usually focuses on the language of statutes, appellate decisions or the work of rival academics, not the standard jury instructions used by trial courts.”); *infra* Parts II, IV.E.

10. The previous failure to ask this sort of question about jurors is puzzling. After all, scholars have offered rich accounts of how *judges* resolve ambiguous legal directives, both in general (e.g., general theories of adjudication like those espoused by legal realists and formalists), and more specifically (e.g., economic analyses of tort law doctrines or statutory

The results are surprising. This Article uncovers important divergences between, on the one hand, what courts, commentators, and code drafters *think* belief-state terms do and should mean, and, on the other hand, what jurors are *told* that they mean. The four main divergences can be summarized as follows.<sup>11</sup> First, the distinction between knowledge and recklessness, as defined in popular jury instructions and applied in practice, is nearly nonexistent. As a result, jurors likely “over”-ascribe knowledge to criminal defendants, and even where they do not, significant sentencing

---

interpretation). Granted, jurors, unlike judges, leave no written explanation of their decision-making. But much of the judicial decision-making literature adopts an “external” perspective that seeks to explain judicial decision-making without recourse to the explanations judges themselves give. See Charles L. Barzun, *Inside-Out: Beyond the Internal/External Distinction in Legal Scholarship*, 101 VA. L. REV. 1203, 1235-36 (2015). For speculation as to some reasons for the previous neglect, see *infra* Section IV.E.

11. Before summarizing them, it may help to provide a more concrete example of the sorts of ambiguities that pervade current mens rea jury instructions. Consider the concept of “knowledge” at work in a recent Supreme Court case, *Rosemond v. United States*, 134 S. Ct. 1240 (2014). Rosemond’s partner brought a gun to a marijuana deal Rosemond helped orchestrate. *Id.* at 1243. At Rosemond’s trial for aiding and abetting an armed drug deal, the prosecution needed to prove that Rosemond “knew” in advance that his partner would show up armed, and that Rosemond “knowingly” participated in an armed drug deal. *Id.* at 1244, 1250. Rosemond alleged he did not know his partner would be armed. *Id.* at 1246. (The aiding and abetting instructions in *Rosemond*, like many federal instructions, did not define knowledge, see Jury Instructions at 46, *United States v. Rosemond*, No. 2:07-CR-886 (D. Utah 2011); *infra* Section II.B, but even MPC-based definitions of knowledge leave open the ambiguities highlighted below, see *infra* Section II.A.)

What might it mean to find that Rosemond “knew” that P, where “P” means “my partner will be armed”? First, and least controversially, for Rosemond to have “known” that P, P must have turned out to be *true*—i.e., Rosemond’s partner must actually bring a gun. Second, it seems likely that Rosemond must have *believed* that P. But what sort of belief, if any, was necessary? Must Rosemond have had an “occurrent” belief (i.e., conscious consideration of P at the relevant time), or instead merely a “dispositional” belief (i.e., some sort of latent acknowledgment that P, such that Rosemond, if asked at the relevant time, would have agreed that P even though he hadn’t been thinking about it)? And would it suffice for Rosemond to have in some sense intellectually acknowledged that P but yet act as if “deep down” he believed or accepted Not-P, perhaps due to “wishful thinking” or some other conative thought process? Third, Rosemond might have to have some higher degree of *subjective certainty* about P than the mere belief that P is more likely true than Not-P. If so, how much certainty must he have? Eighty percent? Ninety percent? Finally, a few overarching questions: does the degree of morality or immorality of the actor’s P-relevant action, in Rosemond’s case or in other cases, impact the type of belief that P, or the degree of certainty that P, that suffices for knowledge that P? And can the requisite type of belief or degree of certainty depend on pragmatic features of the agent’s situation—for instance, how much time Rosemond had available to form an opinion about P? For answers to these questions, see *infra* Section III.B.

disparities hinge on an unreliable distinction. Second, the legal concept of “willful ignorance” is often a form of knowledge. Where it is not, the two are distinguishable largely because of a difference in the actor’s practical circumstances, rather than a difference in the actor’s mental state. Third, important concepts like “belief,” “awareness,” and “conscious disregard” signify different kinds of mental states in different circumstances. As a result, jurors’ belief-state ascriptions systematically track features of the case typically thought irrelevant to mental-state determination, ultimately leading to mistaken convictions and acquittals. Fourth, and more generally, modern mens rea terms are more context-dependent and evaluative than is typically thought. This calls into question the conventional wisdom that modern criminal codes like the MPC employ less evaluative, and more purely descriptive, mens rea concepts than did the common law. It also has important implications for various normative proposals premised on that conventional wisdom.

The Article proceeds as follows. Part II addresses how MPC and non-MPC jurisdictions use and define belief state terms such as “knowledge,” “belief,” “conscious disregard,” and “failure to perceive.” It focuses on actual belief-state jury instructions, as well as commentary from scholars, courts, and code drafters concerning the meaning of those belief-state terms. It shows that jury instructions provide little explanation of belief states, leaving a large role for jurors’ pre-existing understandings of the terms in question. Jurors interpret ambiguous directions in a way that coincides with their natural language-based understanding of the terms at issue. What that understanding is, and how it is employed in response to circumstantial evidence of mental state, is an empirical question.

Part III draws on recent empirical research at the intersection of philosophy and psychology—falling under the banner of “Experimental Epistemology”—to provide a descriptive account of how laypeople interpret and ascribe “belief,” “knowledge,” and related concepts. This literature sheds light on how laypeople resolve the sorts of ambiguities noted in the previous Part’s discussion of common jury instructions. The resulting picture of the various mens rea terms differs in important respects from conventional legal and philosophical accounts of what those terms mean.

Part IV examines the implications of the empirical research for the criminal law’s treatment of belief states. It shows that several assumptions underlying courts’, commentators’, and perhaps legislatures’ mens rea analyses are flawed, resulting in misguided statutory schemes, misleading jury instructions, and unintended effects on case outcomes. For each



divergence between mens rea theory and practice this Part highlights, it also explains why the divergence matters and discusses potential ways to address it. After discussing the four main divergences listed above, Section IV.E considers several objections to the Article's focus on jury instructions. It argues that these objections are unwarranted, and that scholarly scrutiny of mens rea jury instructions is particularly important because current institutional mechanisms are ill-designed to flag, let alone remedy, systematically problematic mens rea instructions.

Part V concludes.

## *II. Belief States in Criminal Law*

This Part examines what jurors are told about the legal definitions of belief states, as well as how legal scholars, legislators, model code drafters, and judges describe and interpret various belief-state terms. It first addresses instructions in MPC jurisdictions and then more briefly examines federal instructions as a particularly important example of non-MPC mens rea analysis.

### *A. The Model Penal Code*

The MPC employs four main mens rea categories, listed here in order of decreasing culpability: (1) purposely, (2) knowingly, (3) recklessly, and (4) negligently.<sup>12</sup> Each entails a minimum requisite belief state, and each leaves open significant questions concerning what constitutes that belief state.

---

12. MPC § 2.02(2). Each state further up the culpability chain suffices to establish those lower on the chain. *Id.* § 2.02(5). The MPC adds a fifth mens rea in the context of homicide, labeled “extreme indifference.” *Id.* § 210.2(1)(b). Some specific crimes and affirmative defenses delineated in the MPC involve yet other mental states including, for example, “belief.” *See, e.g., id.* § 3.04(2)(c) (“[A] person employing protective force may estimate the necessity thereof under the circumstances as he believes them to be when the force is used”); *id.* § 3.05(1)(b) (providing that use of force in defense of another person is to be evaluated on the basis of “the circumstances as the actor believes them to be”); *id.* § 5.01(1)(c) (defining the elements of criminal attempt); *id.* § 223.6(1) (offense of receiving stolen property requires that one “purposely receives, retains, or disposes of movable property of another knowing that it has been stolen, or believing that it has probably been stolen”); *id.* § 241.1 (perjury); *id.* § 3.02(1) (affirmative defense available where actor “believe[d]” conduct was necessary to avoid greater harm). Non-MPC jurisdictions similarly utilize “belief” as a mental state distinct from knowledge, *see, e.g.,* 18 U.S.C. § 1621(2) (2012) (perjury), and jurisdictions that typically follow the MPC at times invoke “belief” where the MPC does not, *see, e.g.,* 27 MD. CODE ANN., Criminal Law § 7-101(b) (LexisNexis 2012) (deception); TEX. PENAL CODE ANN. § 31.0(1)(A) (West Supp. 2014) (deception); WASH. REV. CODE ANN. § 9A.72.110(2) (LexisNexis 2014) (intimidating former witnesses).

Since “purposely” figures less centrally in the MPC,<sup>13</sup> I focus on the other three MPC mental states and leave “purposely” for another day.<sup>14</sup>

### *I. Knowingly*

MPC-based jury instructions concerning knowledge typically provide several definitions. Knowledge of an inculpatory proposition, or knowledge that “P,”<sup>15</sup> is defined as (1) “awareness” that P, or (2) “awareness that it is practically certain” that P, or (3) “awareness” of a “high probability” that P, coupled with a lack of “actual belie[f]” that Not-P.<sup>16</sup> Virtually all MPC-based knowledge instructions contain both definitions (1) and (2), and many include definition (3), presenting all three definitions as nonexclusive alternative bases for finding that the defendant acted knowingly.<sup>17</sup> These definitions contain two central ambiguities, addressed below.

---

13. Simons, *Rethinking*, *supra* note 1, at 470-71 (“[T]he concept of purpose is surprisingly unimportant: although the [MPC] distinguishes between purpose and knowledge in the definitional section, it only rarely distinguishes between them in the sections specifying requirements for individual offenses.”).

14. To be sure, purpose does entail a particular belief state, both as a conceptual matter and as a matter of legal definition under the MPC. See LARRY ALEXANDER & KIMBERLY KESSLER FERZAN, *CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW* 35 (2009) (“[I]n order to act with criminal purpose, the actor must believe that his conduct increases the risk of harm.”); MPC §§ 2.02(a)(ii) (to act “purposely” with respect to an attendant circumstance element, the defendant must be “aware of the existence of such circumstances or he believes or hopes that they exist”); *id.* § 2.02 cmt., at 229-41 (“[T]he Code draws a narrow distinction between acting purposely and knowingly . . . . Knowledge that the requisite external circumstances exist is a common element in both conceptions.”). That said, and somewhat mysteriously, the MPC’s definition of purpose as to a *result* element does *not* affirmatively require that the defendant possess any type of belief that the action at issue would increase the result’s probability. See *id.* § 2.02(2)(a)(i).

15. Throughout this Article, “P” will represent an inculpatory proposition, and “Not-P” will represent its opposite. So, for example, if “P” represents the proposition, “my partner will be armed,” “Not-P” represents the proposition, “my partner will not be armed.”

16. See Simons, *Should the Model*, *supra* note 7, at 182 n.9 (noting that “the Code’s definition of knowledge as to a circumstance must be derived from two sources: the basic definition, ‘aware that such circumstances exist,’ in § 2.02(2)(b)(i); and ‘aware of a high probability of its existence,’ in § 2.02(7),” and that “[t]he latter phrase presumably controls the former, since it is a weaker requirement.”).

17. See Robbins, *supra* note 8, at 226 n.3 (“Although the Code states that this provision is designed to combat the problem of deliberate ignorance, this statement is contained only in a comment. . . . Section 2.02(7), which contains the definition, does not place this restriction on its use.”) (citation omitted); *United States v. Jewell*, 532 F.2d 697, 700 (9th Cir. 1976) (en banc) (reasoning, in the context of willful ignorance, that “[t]o act ‘knowingly’ . . . is not necessarily to act only with positive knowledge, but also to act with

One ambiguity concerns what constitutes “awareness” and “actual belief.” To be “aware” that P, or to “actually believe” that P, must one be consciously considering P at the relevant time (sometimes called an “occurrent” belief state<sup>18</sup>), or could one instead merely be disposed to endorse P if asked, even if one is not consciously considering P at the relevant time (a “dispositional” belief state<sup>19</sup>)? Furthermore, must one have some sort of internal conviction or acceptance of P “deep down,” or could one instead merely intellectually acknowledge that P, without any attendant emotion or conation?<sup>20</sup> Neither the MPC’s drafters nor criminal law scholars have provided much in the way of analysis, leading Professor Douglas Husak to describe as “scandalous” the fact that scholars have “not made greater efforts to understand the nature of awareness,” a “pivotal topic” that is “radically under-theorized.”<sup>21</sup>

Professor Husak begins his own analysis of awareness in the context of negligence and recklessness with a “crucial” observation: “awareness (or belief) need not be *occurrent*.”<sup>22</sup> In other words, one can be “aware” that P or “actually believe” that P, in the legally relevant sense, even if one is not consciously considering P at present—so long as one has the right sort of “disposition” toward P such that one would acknowledge its truth if asked about it under normal conditions. Additionally, commentators construe awareness and belief as not requiring any conation or internal conviction.<sup>23</sup>

---

an awareness of the high probability of the existence of the fact in question”); *accord* Global-Tech Appliances, Inc. v. SEB S.A., 131 S. Ct. 2060, 2069 (2011).

18. “Occurrent” belief is conscious endorsement at a given moment. See Wesley Buckwalter & John Turri, In the Thick of Moral Motivation (Aug. 27, 2014) (manuscript at 6), [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2382383](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2382383).

19. *Id.*

20. For a more complete account of these distinctions, see *infra* Section III.B.1.b.

21. Husak, *supra* note 4, at 207; see also Stephen P. Garvey, *What's Wrong with Involuntary Manslaughter?*, 85 TEX. L. REV. 333, 343-44 (2006); Kenneth R. Feinberg, *Toward a New Approach to Proving Culpability: Mens Rea and the Proposed Federal Criminal Code*, 18 AM. CRIM. L. REV. 123, 128 (1980) (contrasting “awareness” with “belief drawn from inference”).

22. Husak, *supra* note 4, at 208; see also Douglas Husak, *Distraction and Negligence*, in PRINCIPLES AND VALUES IN CRIMINAL LAW AND CRIMINAL JUSTICE: ESSAYS IN HONOUR OF ANDREW ASHWORTH 81, 85 (Lucia Zedner & Julian V. Roberts eds., 2012).

23. An exception might be “depraved heart” or “extreme indifference” homicide. See MPC § 210.2(1)(b). Like some earlier common law mens rea concepts, it appears more overtly tied to a failure to deeply come to grips with or accept the risk that one’s action will result in another’s death. For a more traditional interpretation of “depraved heart” and “extreme indifference” homicide, see Kenneth W. Simons, *Statistical Knowledge Deconstructed*, 92 B.U. L. REV. 1, 57-58 (2012) [hereinafter Simons, *Statistical Knowledge*].

That is, the relevant inquiry in all cases concerning criminal knowledge is into the defendant's *cognitive* disposition—his or her intellectual acknowledgment that P.

Another ambiguity concerns the degree of confidence one must have in the truth of P in order to be “practically certain” that P or to be aware of its “high probability.”<sup>24</sup> Commentators agree on two parts of an admittedly incomplete answer. First, the answer is *not* “it depends on the situation.” Rather, whatever the requisite degree of certainty may be, it is contextually invariant; to know that P, one needs X amount of certainty that P, regardless of what P is. Professor Kenneth W. Simons sums up this consensus view nicely when, in contrasting knowledge with the predominant view of recklessness, he writes, “[o]f course, knowledge is indeed an invariant mental state; when it is required, the actor must be aware of a ‘high probability’ or a ‘practical certainty,’ *period*, without regard to any other factors.”<sup>25</sup>

The second point of agreement among commentators is that, whatever this certainty level is, it is above 50%, and probably considerably higher than 50%. It must be *at least* above 50%, commentators reason, because knowledge that P entails belief that P and belief that P entails acknowledgment of P's being more likely true than not.<sup>26</sup> It must be

---

24. Some jurisdictions use slight variations in wording that are subject to the same ambiguities and analysis here provided. *See, e.g.*, TEX. PENAL CODE ANN. § 6.03(b) (West 1994) (defining knowledge as to a result element as requiring that one be “reasonably certain” that one will cause the result); *see also* 720 ILL. COMP. STAT. ANN. § 5/4-5 (West 2014) (defining “knowledge” as conscious awareness of “substantial probability”). Ohio avoids some of the ambiguities by defining knowledge as awareness that a result or circumstance is merely “probabl[e].” *See* 29 OHIO REV. CODE ANN. § 2901.22(B) (Lexis Nexis 2014).

25. Simons, *Should the Model*, *supra* note 7, at 189-90; *see also* Kimberly Kessler Ferzan, *Holistic Culpability*, 28 CARDOZO L. REV. 2523, 2536, 2529-30 (2007); Alan C. Michaels, *Acceptance: The Missing Mental State*, 71 S. CAL. L. REV. 953, 972 (1998) (distinguishing knowledge from other belief states by virtue of its attendant degree of subjective certainty).

26. *See* MPC § 2.02(7) cmt. 9, at 248 (stating that the invocation of “high” probability is designed to deal with deliberate ignorance, which the Commentary describes as involving “the case of the actor who is aware of the *probable* existence of a material fact but does not determine whether it exists or does not exist”) (emphasis added); Garvey, *supra* note 21, at 371 n.22 (“[T]he text of §2.02(7) presupposes that an actor can at the same time believe that the probability that p exists is high while at the same time believing that p does not exist. But the belief that the probability that p exists is high precludes the belief that p does not exist, or at least precludes holding that belief rationally.”); Jonathan L. Marcus, Note, *Model Penal Code Section 2.02(7) and Willful Blindness*, 102 YALE L.J. 2231, 2255 (1993) (“[I]t is

significantly *above* 50% because (a) “[o]ne would not normally say that one knows something unless one feels fairly certain of it in one’s mind,”<sup>27</sup> and (b) the MPC’s distinction between recklessness and knowledge is premised on there being a significant difference between awareness of a merely “substantial” probability (recklessness) and awareness of a “high” probability (knowledge).<sup>28</sup> The MPC’s drafters apparently considered “high” probability to denote something significantly above 50%, as evidenced by their distinguishing, in the Commentaries, the MPC’s “high probability” form of knowledge from Ohio’s “more expansive” code, which allows for satisfaction of the knowledge requirement where a result or circumstance is merely “probable.”<sup>29</sup>

## 2. *Recklessly and Negligently*

MPC-based jury instructions defining recklessness and negligence use similar structure and phrasing. One is *reckless* if one (a) “consciously disregards” (b) a “substantial and unjustifiable risk,” where (c) “considering the nature and purpose of the actor’s conduct and the circumstances known

---

difficult to imagine how one can simultaneously be aware of a high probability that a fact exists yet believe that it does not exist.”); Husak & Callender, *supra* note 8, at 51 (stating that a drug courier aware of a 33% risk that his suitcase contains narcotics could not believe that there is a “high” probability that his suitcase contains drugs for purposes of MPC § 2.02(7)). *But see Global-Tech*, 131 S. Ct. at 2073 (Kennedy, J., dissenting) (“One can believe that there is a ‘high probability’ that acts might infringe a patent but nonetheless conclude they do not infringe.”).

27. Robin Charlow, *Wilful Ignorance and Criminal Culpability*, 70 TEX. L. REV. 1351, 1373 & n.105 (1992) (“What is important is to indicate that the level of belief is exceptionally high when knowledge is involved.”); see WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR., *CRIMINAL LAW* 239 (2d ed. 1986) (stating that “knowledge” requires at least “a consciousness of almost-certainty”); *United States v. Golamb*, 811 F.2d 787, 792 (2d Cir. 1987) (“Knowledge and belief are very different mental states; knowledge implies a much higher degree of certainty.”). Going further still, Professor Robbins writes that “criminal knowledge requires certainty and a corresponding absence of doubt,” so that the MPC’s definition of knowledge, in allowing for a probability assessment to count as knowledge despite the fact that probabilities “impl[y] an absence of certainty,” fails to track the concept of “criminal knowledge.” Robbins, *supra* note 8, at 222 & n.208.

28. *See, e.g.*, Marcus, *supra* note 26, at 2240 (“[H]igh probability’ entails well over a 51% chance of harm.”); Charlow, *supra* note 27, at 1382 (“[I]n order to ‘know’ one must be aware of the certainty or near certainty of a fact, and in order to be ‘reckless’ one must be aware of, at most, the substantial probability of a fact.”); Simons, *Rethinking*, *supra* note 1, at 474 (“Criminal law distinguishes recklessness from knowledge according to a single factor: whether the actor believed that the risk was merely ‘substantial’ (recklessness) or instead ‘highly probable’ (knowledge).”).

29. *See* MPC § 2.02 cmt. 9, n.43.

to him,” such disregard “involves a gross deviation” from reasonable standards of conduct.<sup>30</sup> One is *negligent* if one (a) “should be aware of” (b) a “substantial and unjustifiable risk,” where (c) “considering the nature and purpose of his conduct and the circumstances known to him,” the actor’s “failure to perceive” the risk involves a gross deviation from reasonable standards of conduct.<sup>31</sup> This Section notes three sources of ambiguity, and their resolution by commentators, below.

The first source of ambiguity concerns recklessness’ requirement that the defendant “consciously disregard” a risk, and negligence’s requirement that the defendant “fail[] to perceive” a risk.<sup>32</sup> Both requirements contain similar ambiguities to those addressed above with respect to “awareness.” For starters, does “conscious disregard” or a “failure to perceive” imply an occurrent state, as opposed to a dispositional one? Scholars are divided.<sup>33</sup> And must “conscious disregard” involve some sort of inner conviction, or does it require merely a cognitive, intellectual acknowledgment that the risk exists? Scholars seem to construe it as requiring only the latter, though without explicitly addressing the distinction. Finally, in order to have “consciously disregard[ed]” a substantial and unjustifiable risk, how certain must the defendant have been that the risk was substantial and/or unjustifiable? Scholars appear not to have addressed the issue.<sup>34</sup>

Second, both recklessness and negligence instructions require consideration of “the circumstances known to” the defendant.<sup>35</sup> Does “knowledge” in this context mean the same thing it means elsewhere in the MPC? Some scholars construe the term differently here, treating it as if it

---

30. *Id.* § 2.02(2)(c).

31. *Id.* § 2.02(2)(d).

32. *Id.* §§ 2.02(2)(c), (d).

33. Garvey, *supra* note 21, at 344 n.55 (collecting sources).

34. Scholars differ over the distinct but related question of whether the reckless actor need recognize both the risk’s substantiality and its unjustifiability, or instead merely one or the other. *See, e.g.*, Claire Finkelstein, *Responsibility for Unintended Consequences*, 2 OHIO ST. J. CRIM. L. 579, 594-95 (2005) (arguing that the reckless actor must be aware of the risk’s substantiality); Joshua Dressler, *Does One Mens Rea Fit All?: Thoughts on Alexander’s Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 955 (2000) (arguing that the reckless actor need be aware only of the risk’s unjustifiability); David M. Treiman, *Recklessness and the Model Penal Code*, 9 AM. J. CRIM. L. 281, 362 (1981) (arguing that “the actor need not be aware that the risk is unjustifiable”). The question I raise here is not what the object of the reckless actor’s awareness must be—e.g., the risk’s substantiality, its unjustifiability, or both—but rather, what it means to be “aware” of something in the relevant sense.

35. MPC §§ 2.02(2)(c)-(d).

means merely “belief.”<sup>36</sup> Others argue that here, as elsewhere, knowledge requires more than mere belief, pointing out that the MPC drafters, in numerous spots where they intended to invoke mere belief, use the word “belief” and not “knowledge.”<sup>37</sup>

Third, both recklessness and negligence instructions depend on whether the risk at issue was “substantial and unjustifiable.”<sup>38</sup> Consider the “substantial” requirement. How much risk counts as “substantial” risk? Some scholars treat substantiality as wholly dependent upon unjustifiability, so that there is no substantiality requirement, only unjustifiability.<sup>39</sup> Other scholars, hesitant to read the substantiality requirement out of the Code, favor an interpretation according to which a “substantial” risk is one that surpasses an unspecified but contextually invariant likelihood threshold—though one that is far below 50%.<sup>40</sup> Now consider the “unjustifiable” requirement. How risky must something be in order to constitute an “unjustifiable” risk? All agree that the answer is, roughly, “it depends on the situation”—i.e., on how bad it would be if the risk came to fruition and on what reasons the actor has for engaging in the risky activity.<sup>41</sup> In other words, the “unjustifiability” requirement for recklessness and negligence, along with the invocation of a “gross

---

36. See Johnson, *supra* note 7, at 514 n.30 (collecting sources).

37. See *id.* at 515-27 (arguing that the MPC’s historical background supports the contention that its invocation of knowledge, as opposed to mere belief, in section 2.02(2)(c)-(d) was intentional); Eric A. Johnson, *Is the Idea of Objective Probability Incoherent?*, 29 LAW & PHIL. 419, 428-29 (2010) (arguing that the MPC’s use of both the phrase “circumstances known to [him]” and the phrase “circumstances as he believes them to be” reflects the drafters’ awareness of the distinction between the two); see also examples cited *supra* note 12.

38. MPC §§ 2.02(2)(c)-(d).

39. See LAFAVE & SCOTT, *supra* note 26, at 239 (“[I]f there is no social utility in doing what he is doing, one might be reckless though the chances of harm are something less than 1%.”); see also Dressler, *supra* note 34, at 959 (arguing that “substantial” should be read to modify “unjustifiable,” such that reckless actors must believe they are taking a “substantially unjustified” risk, and not necessarily that the risk is substantial); ALEXANDER & FERZAN, *supra* note 14, at 25 (arguing that “the ‘substantiality’ prong of the definition should be eliminated”).

40. See Simons, *Should the Model*, *supra* note 7, at 190 (“If this threshold view is correct, then, in order for an actor to be reckless as to causing bodily injury to another, he must be aware of at least (say) a 5% risk that his blow will injure the victim.”); Treiman, *supra* note 34, at 337-38 (“The second function that the requirement of substantiality might serve is as an exclusion of de minimis violations of the law.”).

41. See, e.g., GEORGE P. FLETCHER, BASIC CONCEPTS OF CRIMINAL LAW 115 (1998); LAFAVE & SCOTT, *supra* note 27, at 239.

deviation” from “reasonable” standards of conduct, makes these mental-state concepts contextually “variant” and normatively “evaluative.” They are thus unlike knowledge, which is supposed to be contextually “invariant” and “purely descriptive.”<sup>42</sup>

*B. The Federal Criminal Code*

The federal criminal code employs a far greater array of mental state terms than does the MPC,<sup>43</sup> and these terms are typically left undefined in the statutes in which they are found. Moreover, the same terms are defined differently depending on which substantive crime is at issue, which federal circuit one is in, and which judge within a given district is crafting the jury’s instructions. Nonetheless, a broad overview is possible. It is also worthwhile: the federal courts are a particularly important example of a non-MPC jurisdiction, and federal judges possess greater discretion than do judges in MPC jurisdictions in both (a) defining mental state terms invoked in statutes, and (b) choosing which mental state terms to employ where the statute is silent.<sup>44</sup>

---

42. See, e.g., SAMUEL PILLSBURY, JUDGING EVIL: RETHINKING THE LAW OF MURDER AND MANSLAUGHTER 83-85 (1998); Simons, *Should the Model*, *supra* note 7, at 199; Kenneth W. Simons, *Understanding the Topography of Moral and Criminal Law Norms*, in PHILOSOPHICAL FOUNDATIONS OF CRIMINAL LAW 228, 248 n.51 (R.A. Duff & Stuart P. Green eds., 2011) [hereinafter Simons, *Understanding the Topography*]; 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, 725; Ferzan, *supra* note 24, at 2536, 2529-30; Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 931, 940 (2000).

43. See S. REP. NO. 95-605, pt. 1, at 55 (1977) (“The National Commission’s consultant [on Reform of the Federal Criminal Law] . . . identified 78 different [mens rea] terms used in present law.”); Feinberg, *supra* note 21, at 125. Moreover, where the federal code’s provisions fail to specify a mens rea, federal courts typically supply one as a matter of interpretation. See Darryl K. Brown, *Federal Mens Rea Interpretation and the Limits of Culpability’s Relevance*, 75 LAW & CONTEMP. PROBS., no. 2, 2012, at 109, 113.

44. See MPC § 2.02(3); Brown, *supra* note 43, at 113; *United States v. Bailey*, 100 S. Ct. 624, 630 (1980) (“Few areas of criminal law pose more difficulty than the proper definition of the *mens rea* required for any particular crime.”). Recent scholarship contains a surprising dearth of systematic attention to federal mens rea terms, particularly given the increasing importance of, and scrutiny of, the federal criminal code more generally. See Robert H. Joost, *Federal Criminal Code Reform: Is It Possible?*, 1 BUFF. CRIM. L. REV. 195, 195 (1997).



There are three primary ways federal criminal jury instructions address the concept of knowledge.<sup>45</sup> First, some instructions leave undefined what it means to “know” that P. For example, jury instructions in the Eighth Circuit tend not to define the term, since, according to the Eighth Circuit Court of Appeals, its meaning is “a matter of common knowledge.”<sup>46</sup> Even outside the Eighth Circuit the term is often undefined,<sup>47</sup> especially, though certainly not exclusively, where it is invoked indirectly, as part of the definition of another mens rea term such as “intentionally” or “willfully.”<sup>48</sup> Second, many federal instructions employ the MPC’s formulation of knowledge, defining it as “awareness of a high probability” that P coupled with lack of “actual belief” that Not-P.<sup>49</sup> Such instructions are often given in cases where the prosecution presents some minimal evidence of willful

---

45. *But see* S. REP. NO. 96-553, at 60 (1980) (noting appellate courts’ use of five different meanings of “knowledge” in the federal criminal code). The Senate Report explains,

[knowledge] has been defined in terms of awareness; in terms of a defendant’s inference from the circumstances or belief that something is probably true; in terms of a defendant’s awareness of a “high probability” that a circumstance exists; in terms of intentional or purposeful or “studied ignorance” as to the existence of a fact; and in terms of “gross indifference to” or “willful neglect of” a duty in respect to ascertainment of particular facts.

*Id.*

46. *United States v. Evans*, 431 F.3d 342, 347-48 (8th Cir. 2005); *United States v. Brown*, 33 F.3d 1014, 1017 (8th Cir. 1994); *see also* MANUAL OF MODEL CRIMINAL JURY INSTRUCTIONS FOR THE DISTRICT COURTS OF THE EIGHTH CIRCUIT § 7.03 (Judicial Comm. on Model Jury Instructions for the Eight Circuit 2013).

47. *See, e.g.*, *United States v. Aguilar*, 80 F.3d 329, 331 (9th Cir. 1996); PATTERN CRIMINAL JURY INSTRUCTIONS § 2.06 (Comm. on Pattern Criminal Jury Instructions Dist. Judges Ass’n Sixth Circuit 2013). Sometimes the term “knowingly” is defined in jury instructions simply as “with knowledge.” *See, e.g.*, *United States v. Guzman*, 781 F.2d 428, 431 (5th Cir. 1986).

48. In *Rosemond*, for example, as is common with aiding and abetting instructions, the instruction simply stated that the jury must determine whether the defendant “knew” that his partner was, or would be, armed. Jury Instructions at 46, *United States v. Rosemond*, No. 2:07-CR-886 (D. Utah 2011). Although the Supreme Court’s decision with respect to those instructions spoke of knowledge as “awareness,” “full awareness,” or “full knowledge,” the Court gave no indication that the instructions erred in not defining the term itself as it appeared in the count for aiding and abetting. *See Rosemond v. United States*, 134 S. Ct. 1240, 1248-50 (2014).

49. *See, e.g.*, *Global-Tech Appliances, Inc. v. SEB S.A.*, 131 S. Ct. 2060, 2069 (2011); *Leary v. United States*, 395 U.S. 6, 46 (1969) (using MPC § 2.02(7)’s definition of “knowledge” where neither statutory definition nor relevant legislative history was available). As these and many other cases indicate, non-MPC jurisdictions such as the federal courts are heavily influenced by the MPC.

ignorance.<sup>50</sup> They typically add a requirement that the defendant purposefully avoided acquiring additional certainty of P, though that purposeful avoidance may simply be a failure to investigate when it would have been easy to do so.<sup>51</sup> Third, federal jury instructions often state that “[t]he word ‘knowingly’ . . . means that the act was done voluntarily and intentionally and not because of mistake or accident,” or some close variation.<sup>52</sup> Such instructions use a much more minimal conception of knowledge than that provided in MPC-based instructions and other federal instructions.<sup>53</sup>

As for belief states other than “knowledge,” federal instructions frequently contain ambiguities similar to those previously noted with respect to recklessness and negligence under the MPC.<sup>54</sup> For example, they employ terms like “conscious avoidance” and “awareness” that are

---

50. *But see* 1 LEONARD B. SAND ET AL., MODERN FEDERAL JURY INSTRUCTIONS § 3A.01, at 11 (2011) (stating that “the giving of this instruction” absent evidence of willful ignorance in a federal circuit that requires such evidence “is subject to the harmless error rule, and has been excused on numerous occasions”) (collecting cases).

51. *But see* Michaels, *supra* note 25, at 988-89 (“In practice, the purposeful avoidance requirement adds practically nothing” since “liability may be predicated on omission to learn: The defendant need not close his eyes to be liable; he can be convicted for failure to investigate.”). Indeed, Judge Posner’s opinion in *United States v. Giovannetti*, often cited as a more strenuous assertion of the need for evidence of “deliberate effort to avoid guilty knowledge,” notes that such effort “can be a mental, as well as a physical, effort—a cutting off of one’s normal curiosity by an effort of will.” 919 F.2d 1223, 1229 (7th Cir. 1990).

52. *See, e.g.*, PATTERN JURY INSTRUCTIONS (CRIMINAL CASES) § 2.13 (Comm. on Pattern Criminal Jury Instructions First Circuit 1998); ELEVENTH CIRCUIT PATTERN JURY INSTRUCTIONS (CRIMINAL CASES) § 9.1A (Comm. on Pattern Criminal Jury Instructions of the Judicial Council of the Eleventh Circuit 2010). Such instructions, though common, have received little attention from commentators discussing mental state requirements. This is noteworthy in part because, if taken literally, popular versions of the instruction imply that a defendant who was ignorant of P, but who would have taken the same action had she known that P, acted “knowingly” with respect to P, since such a defendant acts “intentionally and voluntarily, and *not because of* ignorance or mistake.” *Cf.* Michaels, *supra* note 25 (arguing that a similar counterfactual-based mental state should be employed in criminal law).

53. Despite their seemingly minimal requirements for knowledge, it bears emphasis that neither these instructions, nor other federal criminal instructions, use the term “believe” as a substitute for “know.” In other words, in the jury instructions that follow the abstract definition of “to act knowingly,” the issue is always presented as whether the defendant “knew that” P, not whether the defendant merely “believed that” P.

54. *See, e.g.*, ALEXANDER & FERZAN, *supra* note 14, at 25 (noting that the MPC recklessness “formulation is substantially the same as the formulations of recklessness in federal and state criminal codes and judicial decisions”).

ambiguous with respect to occurrent versus dispositional belief.<sup>55</sup> Likewise, federal statutes and instructions refer to what the defendant “actually believe[d]” or had “reason[] to believe,” where belief may be construed as a “deep down” emotional or conative acceptance or as merely intellectual acknowledgment.<sup>56</sup>

### *III. An Empirically Informed Account of Belief State Ascription*

The previous Section gave an overview of the many places where criminal law invokes belief states. It also highlighted ambiguities in jury instructions with respect to belief states, noting previous attempts to resolve those ambiguities and thus to explain precisely what factual findings are required to satisfy various mens rea requirements. This Section temporarily sets aside commentators’ attempts to make sense of those legally salient concepts and examines instead how laypeople, in response to circumstantial evidence, interpret and ascribe belief states. In doing so, it draws from experimental epistemology, an area of empirical research previously overlooked in the legal literature.

#### *A. Introducing Experimental Epistemology*

Philosophers (and legal theorists) often appeal to supposedly widely shared intuitions about what does and does not constitute a given belief state.<sup>57</sup> Experimental epistemologists investigate empirically whether laypeople actually share the intuitions to which “armchair” philosophers appeal, in the hopes that philosophical analyses of belief-state concepts can

---

55. See, e.g., *United States v. Lange*, 528 F.2d 1280, 1288 (5th Cir. 1976) (stating that “conscious purpose to avoid learning the truth” plus “reckless disregard of the truth” suffices for knowledge that a statement was false for purposes of 18 U.S.C. § 1001).

56. See, e.g., 18 U.S.C. § 2339 (2012) (criminalizing the harboring or concealing of those the defendant “knows, or has reasonable grounds to believe,” have committed or are about to commit certain offenses); *Leary v. United States*, 395 U.S. 6, 46 n.93 (1969) (utilizing the MPC’s definition of knowledge, which requires an absence of “actual belief” that Not-P); *United States v. Gansman*, 657 F.3d 85, 94 (2d Cir. 2011) (upholding a “conscious avoidance” instruction employing the “actual belief” language from the MPC).

57. For examples from criminal theory, see sources cited *supra* notes 7-8. For a recent example from evidence theory, see David Enoch & Talia Fisher, *Sense and “Sensitivity”*: *Epistemic and Instrumental Approaches to Statistical Evidence*, 67 STAN. L. REV. 557, 573-74 (2015). For critical discussion of philosophers’ use of intuition, along with examples from the philosophical literature, see, for example, Jonathan M. Weinberg et al., *Are Philosophers Expert Intuiters?*, 23 PHIL. PSYCHOL. 331 (2010).

be made to more accurately account for lay-usage.<sup>58</sup> Whatever the implications for philosophical analyses of the concepts in question, experimental epistemologists' findings concerning how laypeople ascribe belief states help fill a crucial gap in our understanding of legally relevant belief states.

Before turning to belief states specifically, this Section begins with an introduction to one of the central findings in experimental philosophy: the "Side-Effect Effect." Although the Side-Effect Effect concerns desire-state and causal-influence ascription rather than belief-state ascription, understanding it will help in understanding its close relative, the *Epistemic Side-Effect Effect*, which concerns belief-state ascription.

### 1. *The Side-Effect Effect*

In his seminal 2003 study of desire-state ascription (more specifically, intentionality), Professor Joshua Knobe presented subjects with the following vignette (with two variations indicated in *italics*):

The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. We are sure that it will help us increase profits, and it will also *help/harm* the environment." The chairman of the board answered, "I don't care at all about *helping/harming* the environment. I just want to make as much profit as I can. Let's start the new program." They started the new program. Sure enough, the environment was *helped/harmed*.<sup>59</sup>

Subjects in the "help" condition were asked, "Did the chairman help the environment intentionally?" A large majority said "no."<sup>60</sup> Subjects in the "harm" condition were asked whether the chairman harmed the environment intentionally. A large majority said "yes."<sup>61</sup> This phenomenon—laypeople's ascription of intentionality to agents bringing about counter-normative side effects but not to agents bringing about norm-

---

58. For an introduction to this and other research projects in experimental epistemology, see *ADVANCES IN EXPERIMENTAL EPISTEMOLOGY* (James R. Beebe ed., 2014).

59. Joshua Knobe, *Intentional Action and Side Effects in Ordinary Language*, 63 *ANALYSIS* 190 (2003).

60. *Id.* at 195 (77% of subjects).

61. *Id.* (82% of subjects).

compliant side effects<sup>62</sup>—is known as the “Knobe Effect” or the “Side-Effect Effect” (SEE).

Since Knobe’s original study, a vast literature has replicated the same basic asymmetry in a broad range of cases.<sup>63</sup> These studies altered, among other things, the characters, story structure, outcomes, and the properties that study participants were asked to ascribe.<sup>64</sup> It turns out that the SEE is observable not just where an actor produces a bad *side-effect* (whether or not he is indifferent to the side-effect, as the chairman purports to be, or instead actively dislikes it), but also where an actor knowingly causes something bad to happen as a *means to an end* (whether or not the end is good or bad).<sup>65</sup> And people’s willingness to ascribe intentionality in bad-outcome cases persists even where the agent knew only of a *very small risk* that the outcome would come about.<sup>66</sup> It also turns out that the asymmetry observed in intentionality ascription applies to a host of *other desire-state ascriptions* such as being “in favor of” the good or bad outcome.<sup>67</sup> And even attributions of *causal influence*—i.e., the degree of causal responsibility attributed to the act in question as opposed to other partial causes—are significantly higher where an agent’s action results in a

---

62. I use the term “norm-compliant” to mean merely “not in violation of a salient norm.” To be norm-compliant, an action or an outcome need not have resulted from any desire or motivation to comply with a norm; it is enough that the action or outcome happens to comply with norms. See Richard Holton, *Norms and the Knobe Effect*, 70 ANALYSIS 1 (2010).

63. See Joshua Knobe, *Person as Scientist, Person as Moralist*, 33 BEHAV. & BRAIN SCI. 315, 317-20 (2010) (reviewing the literature); *id.* at 329-53 (open peer commentary, including reporting of additional experimental results).

64. Additionally, similar results were obtained in tests of Hindi speakers (using the Hindi word for “intentionally”), as well as with young children and those of different gender, education level, religion, and other demographic attributes. See Joshua Knobe & Arudra Burra, *The Folk Concepts of Intention and Intentional Action: A Cross-Cultural Study*, 6 J. COGNITION & CULTURE 113 (2006); Alan M. Leslie et al., *Acting Intentionally and the Side-Effect Effect: Theory of Mind and Moral Judgment*, 17 PSYCHOL. SCI. 421 (2006); Liane Young et al., *Does Emotion Mediate the Relationship Between an Action’s Moral Status and Its Intentional Status? Neuropsychological Evidence*, 6 J. COGNITION & CULTURE 291 (2006).

65. See, e.g., Florian Cova & Hichem Naar, *Side-Effect Effect Without Side Effects: The Pervasive Impact of Moral Considerations on Judgments of Intentionality*, 25 PHIL. PSYCHOL. 837 (2012).

66. See Mueller et al., *supra* note 2, at 860.

67. See Dean Pettit & Joshua Knobe, *The Pervasive Impact of Moral Judgment*, 24 MIND & LANGUAGE 586 (2009).

foreseen negative outcome rather than a foreseen positive one.<sup>68</sup> Finally, it turns out that the asymmetry is not merely a response to immorality on the part of the agent. Instead, the SEE is observable where an action or outcome deviates from *normative expectations*, including where the deviation is from non-moral descriptive norms,<sup>69</sup> and, perhaps most surprisingly, where the action is regarded as morally praiseworthy but deviates from prudential norms (i.e., is against the agent's self-interest).<sup>70</sup>

---

68. Craig Roxborough & Jill Cumby, *Folk Psychological Concepts: Causation*, 22 PHIL. PSYCH. 205 (2009); Hitchcock & Knobe, *supra* note 6, at 602-05; Mark Alicke, *Culpable Control and the Psychology of Blame*, 126 PSYCH. BULL. 556 (2000); Joshua Knobe & Ben Fraser, *Causal Judgment and Moral Judgment: Two Experiments*, in MORAL PSYCHOLOGY 441 (Walter Sinnott-Armstrong ed., 2008); Lawrence M. Solan & Jonathan M. Darley, *Causation, Contribution, and Legal Liability: An Empirical Study*, LAW & CONTEMP. PROBS., Autumn 2001, at 265 (vol. 64, no. 4); cf. Antony Honoré, *Causation in the Law*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2010), <http://plato.stanford.edu/archives/win2010/entries/causation-law/> (echoing a central contention in H.L.A. HART & TONY HONORÉ, CAUSATION IN THE LAW (2d ed. 1985), that in law as in normal conversation, “[t]he most persuasive explanations of an outcome are those that point to a condition that is abnormal or unexpected in the context or to a deliberate action designed to bring the outcome about”).

69. See Mark Alfano et al., *The Centrality of Belief and Reflection in Knobe-Effect Cases: A Unified Account of the Data*, 95 MONIST 264 (2012) (experiments concerning descriptive norms, conventional norms, and aesthetic norms); Joshua Knobe, *Folk Psychology and Folk Morality: Response to Critics*, 24 J. THEORETICAL & PHIL. PSYCHOL. 270 (2004) (aesthetic norms); Joshua Knobe, *Reason Explanation in Folk Psychology*, 31 MIDWEST STUD. PHIL. 90 (2007) [hereinafter Knobe, *Reason Explanation*] (competing legal and moral norms); Joshua Knobe and Gabriel Mendlow, *The Good, the Bad and the Blameworthy: Understanding the Role of Evaluative Reasoning in Folk Psychology*, 24 J. THEORETICAL & PHIL. PSYCHOL. 252 (2004) (prudential norms); Kevin Uttich & Tania Lombrozo, *Norms Inform Mental State Ascriptions: A Rational Explanation for the Side-Effect Effect*, 116 COGNITION 87 (2010) (conventional norms).

70. For example, in one study, subjects were presented the following vignette:

Imagine that Steve and Jason are two friends who are competing against one another in an essay competition. Jason decides to help Steve edit his essay. Ellen, a mutual friend, says, “Don’t you realize that if you help Steve, you will decrease your own chances of winning the competition?” Jason responds, “I know that helping Steve decreases my chances of winning, but I don’t care at all about that. I just want to help my friend!” Sure enough, Steve wins the competition because of Jason’s help.

Subjects were asked (a) how much praise Jason deserves and (b) whether Jason intentionally decreased his own chances of winning. See Thomas Nadelhoffer, *On Praise, Side Effects, and Folk Ascriptions of Intentionality*, 24 J. THEORETICAL & PHIL. PSYCHOL. 196, 209 (2004). Participants considered Jason praiseworthy, while regarding the side effect as intentional at higher rates than in comparable norm-adhering vignettes. *Id.* at 210. For additional examples of the SEE in the cases of morally positive norm-deviation, see Joshua

These findings are all puzzling. Why would the normative implications of the chairman's (or other protagonist's) *action* affect the degree of intention, causal responsibility, etc., that laypeople ascribe to him or her? Scholars have proposed numerous explanations for the SEE that, for present purposes, can be divided into two camps: competence theories and error theories.

First, competence theories posit that laypeople are competently applying the relevant concepts to the vignettes they are presented, but that it turns out concepts like intentionality and causality, properly understood and applied, depend in part on the normative valence of the actor's action.<sup>71</sup> In other words, competence theorists think that the SEE derives from a correct understanding of the concepts of intentionality and causality. Turning to law, some competence theorists have argued that law should reflect laypeople's conceptual understandings, and that law should therefore, more often than it currently does, treat foreknowledge of an outcome as sufficient for a finding that the actor intentionally caused that outcome.<sup>72</sup>

The second group of explanations, error theories, posit that the SEE is a symptom of biased, or "motivated," reasoning<sup>73</sup>—specifically, a desire to blame those who bring about a bad outcome, regardless of mental state.<sup>74</sup> Such error theories are sometimes called "blame first" models, because they posit that an initial impulse to blame leads people to ascribe whatever mental or causal attribute might justify that initial blame impulse.<sup>75</sup>

---

Knobe, *Reason Explanation*, *supra* note 69, at 102, 105-06 (protagonist violating a racial identification law akin to one used in Nazi Germany deemed to have more intentionally brought about side effect than same protagonist *adhering* to the racial identification law); Brian Robinson et al., *Reversing the Side-Effect Effect: The Power of Salient Norms*, 172 PHIL. STUDY. 177 (2015); *cf.* Holton, *supra* note 62, at 2-3.

71. *See, e.g.*, Holton, *supra* note 62, at 4; Solan, *supra* note 2, at 525; Kobick & Knobe, *supra* note 2, at 413.

72. *See, e.g.*, Kobick & Knobe, *supra* note 2 (ambiguous statutory provisions should be construed to reflect laypeople's understandings and should therefore treat knowledge as sufficient for intent); Kobick, *supra* note 2 (same for constitutional law); Malle & Nelson, *supra* note 2 (same for criminal law); Joseph C. Mauro, *Intentional Killing Without Intending to Kill: Knobe's Theory as a Rational Limit on Felony Murder*, 73 LA. L. REV. 1011 (2013) (felony murder rule should track lay-concept of intentionality).

73. *See* Peter H. Ditto et al., *Motivated Moral Reasoning*, in MORAL JUDGMENT AND DECISION MAKING: THE PSYCHOLOGY OF LEARNING AND MOTIVATION 307 (D. M. Bartels et al. eds., 2009); Eileen Braman & Thomas E. Nelson, *Mechanisms of Motivated Reasoning? Analogical Perception in Discrimination Disputes*, 51 AM. J. POL. SCI. 940, 941 (2007).

74. *See, e.g.*, Mark D. Alicke, *Blaming Badly*, 8 J. COGNITION & CULTURE 179 (2008).

75. Bertram F. Malle et al., *Moral, Cognitive, and Social: The Nature of Blame*, in SOCIAL THINKING AND INTERPERSONAL BEHAVIOR 313 (Joseph P. Forgas et al. eds., 2012).

Conversely, competence theories posit that attributions of blame come later in the process, after evaluation of the agent's mental state.<sup>76</sup> Error theorists typically argue that the law should not embrace laypeople's ascription practices.<sup>77</sup> Instead, error theorists think the SEE is evidence of undesirably normative considerations entering into mental-state and causality ascriptions that are supposed to be purely descriptive.

Neither explanation fully accounts for the data. Competence theorists have trouble explaining why otherwise-irrelevant character information—for example, telling subjects that the protagonist was a drug dealer—increases desire-state and causality ascription (i.e., increases the extent to which laypeople find that the agent (a) intended to bring about the bad outcome and (b) was causally responsible for the outcome despite alternative causes).<sup>78</sup> Likewise, error theorists have trouble explaining why the SEE occurs in cases where the agent's norm-deviation is morally positive or morally neutral.<sup>79</sup> They also have trouble substantiating the claim that blame attribution comes *prior* to mental state attribution, rather than the other way around.<sup>80</sup>

Until recently, both explanations, along with all scholarship concerning the SEE, were also premised on a crucial and unfounded assumption—namely, that study participants regarded the protagonists in the help and harm condition as *knowing* that the good or bad outcome would result.<sup>81</sup>

---

76. *Id.*

77. See, e.g., Nadelhoffer, *Bad Acts*, *supra* note 2.

78. See Nadler & McDonnell, *supra* note 2.

79. See Alfano, *supra* note 69, at 277-78 (discussing Nadelhoffer's "distortion theory" and concluding that it "will not suffice" to explain the SEE given the countervailing data); Holton, *supra* note 62, at 2-5; Knobe, *supra* note 63, at 322 ("[T]he experimental results again and again seemed to go against what would have been predicted on the motivational bias view [i.e., an "error theory" view]. At this point, the vast majority of researchers working on these questions have therefore concluded that the motivational bias hypothesis cannot explain the full range of experimental findings and that some other sort of psychological process must be at work here.") (collecting sources); Edouard Machery, *The Folk Concept of Intentional Action: Philosophical and Experimental Issues*, 23 *MIND & LANGUAGE* 165 (2008); Jason Turner, *Folk Intuitions, Asymmetry, and Intentional Side-Effects*, 24 *J. THEORETICAL & PHIL. PSYCHOL.* 214 (2004).

80. See Malle, *supra* note 76, at 323 (concluding that "the key claim of [blame early] models is, at present, not well supported"); Fiery Cushman, *Crime and Punishment: Distinguishing the Roles of Causal and Intentional Analyses in Moral Judgment*, 108 *COGNITION* 353 (2008).

81. See James R. Beebe & Mark Jensen, *Surprising Connections Between Knowledge and Action: The Robustness of the Epistemic Side-Effect Effect*, 25 *PHIL. PSYCHOL.* 689, 709-11 (2012) (providing numerous examples in the literature).



That is, scholars sought to explain why subjects treated actors who “knowingly” brought about worse outcomes as if they had intentionally caused the outcomes. Until recently, however, nobody examined whether study participants were actually ascribing knowledge, or even belief, to the agent in both conditions. Scholars failed to examine whether the differences in desire-state and causality ascription might be attributable to differences in something more fundamental: belief-state ascription. As the next Section explains, that appears to be exactly what was happening.

## 2. *The Epistemic Side-Effect Effect*

It turns out that belief-state ascription reveals the same asymmetry previously noted for ascription of intentionality, causality, and other attributes.<sup>82</sup> The first study to show this gave participants Knobe’s original chairman vignette and asked whether the chairman “knew” that the program would harm (or help) the environment. Respondents were significantly more likely to attribute knowledge to the chairman in the harm case than in the help case.<sup>83</sup>

In subsequent studies, this phenomenon—the “Epistemic Side-Effect Effect” (ESEE)—has been consistently shown in the broad range of settings in which the SEE has been shown.<sup>84</sup> In short, laypeople are significantly more likely to ascribe belief that P and knowledge that P where “P” is that a norm-deviant outcome will result, that a norm-deviant circumstance obtains, or that one’s action is norm-deviant, compared to where P is norm-

---

82. See *id.* at 691; John Turri, *The Problem of ESEE Knowledge*, 1 ERGO 101 (2014); Wesley Buckwalter, *Gettier Made ESEE*, 27 PHIL. PSYCH. 368 (2014) [hereinafter Buckwalter, *Gettier Made ESEE*]; James Beebe & Joseph Shea, *Gettierized Knobe Effects*, 10 EPISTEME 219 (2013); James R. Beebe, *A Knobe Effect for Belief Ascriptions*, 4 REV. PHIL. & PSYCH. 235 (2013) [hereinafter Beebe, *A Knobe Effect*]; Nikolaus Dalbauer, & Andreas Hergovich, *Is What Is Worse More Likely?—The Probabilistic Explanation of the Epistemic Side-Effect Effect*, 4 REV. PHIL. & PSYCH. 639 (2013); James R. Beebe & Wesley Buckwalter, *The Epistemic Side-Effect Effect*, 25 MIND & LANGUAGE 474 (2010); James R. Beebe, *Evaluative Effects on Knowledge Attributions*, in THE BLACKWELL COMPANION TO EXPERIMENTAL PHILOSOPHY (Justin Sytsma & Wesley Buckwalter, eds.) (forthcoming 2016); Alfano, *supra* note 68.

83. Beebe & Buckwalter, *supra* note 82, at 476.

84. See *supra* notes 63-70, 82. To be sure, there are more published SEE experiments than ESEE experiments. After all, the latter effect was discovered relatively recently. Nonetheless, ESEE experiments to date have largely used the basic experimental structure and vignettes that the SEE studies use, and have found results that consistently track the SEE experiment results in these various settings.

*consistent*.<sup>85</sup> While the SEE showed that normative valence impacts desire-state and causality ascription, the ESEE shows that it also impacts belief-state ascription.

The ESEE largely explains the SEE. In the chairman studies, for example, people are more likely to ascribe intentionality, desire, causation, etc., in the “harm” case than in the “help” case *because* they are more likely to ascribe to the chairman *belief* that the new policy will harm the environment, as well as *knowledge* that the policy will harm the environment. Indeed, take away the differential belief and knowledge ascriptions, and the SEE nearly disappears.<sup>86</sup>

But insofar as the ESEE explains the SEE, the question becomes: what explains the ESEE? Why would the question of whether or not one believes that P or knows that P depend on the normative valence of one’s P-relevant *action*? This seems contrary to just about any philosophical or legal understanding of “knowledge” and “belief,” which are traditionally considered purely descriptive, non-evaluative mental state concepts.<sup>87</sup>

As with the SEE, the ESEE may be explained using a competence theory or an error theory. A competence theory would say that laypeople are competently applying the relevant concepts to the vignettes they are

---

85. Indeed, the same asymmetry appears even where: (a) subjects in the “harm” condition are told that the chairman was notified only that there was a “slight chance” that the new program would harm the environment, and (b) subjects in the “help” condition are told that the chairman was notified that there was a “very strong chance” that the new program would help the environment. See Beebe & Jensen, *supra* note 81, at 707; Dalbauer, *supra* note 82.

86. In situations where the chairman receives no information about the environmental impact of the plan one way or another (i.e., nobody tells him it will help or harm it), the SEE diminishes greatly. Sandra Pellizzoni et al., *Beliefs and Moral Valence Affect Intentionality Attributions: The Case of Side Effects*, 1 REV. PHIL. & PSYCHOL. 201, 201 (2010). And when subjects are explicitly told that the chairman in the harm condition believes the policy will not harm the environment, and the chairman in the help condition believes the policy will not help the environment, the SEE disappears completely. *Id.*

87. See sources cited *supra* note 42, *infra* notes 208-14 and accompanying text; N. Angel Pinillos & Shawn Simpson, *Experimental Evidence Supporting Anti-Intellectualism About Knowledge*, in ADVANCES IN EXPERIMENTAL EPISTEMOLOGY, *supra* note 58, at 9. But see John Turri & Wesley Buckwalter, *Descartes’ Schism, Locke’s Reunion: Completing the Pragmatic Turn in Epistemology*, AM. PHILOSOPHICAL Q. (forthcoming 2016) (tracing to John Locke the roots of philosophical accounts of knowledge as linked to “actionability”). To be more precise, the traditional conception of “knowledge” is evaluative insofar as it entails “justification,” which is a normative concept. But the traditional notion of justification is *epistemically* evaluative; it is not supposed to concern evaluative considerations of the sort found in ESEE cases.

presented, but that it turns out concepts like “belief” and “knowledge,” properly understood and applied, depend in part on the normative valence of the actor’s action.<sup>88</sup> Turning to law, a competence theorist could argue that law should, for example, allow smaller degrees of certainty to constitute “knowledge” where the agent’s action was especially counter-normative.<sup>89</sup>

Alternatively, one might be tempted to endorse an error-based explanation of the ESEE. Perhaps laypeople see a bad outcome, have an impulse to blame somebody, and satisfy that impulse by ascribing whatever mental state seems to inculcate the actor who brought the outcome about.<sup>90</sup> Turning to law, an error theorist could argue that the ESEE is evidence of undesirably normative considerations entering into what should be purely descriptive belief-state ascriptions. As a result, the error theorist might propose means of “de-biasing” jurors’ belief-state ascriptions so that they are unaffected by how good or bad the defendant’s actions were.

Although an error theory may be a good partial explanation for the ESEE, it fails to account for much of the data.<sup>91</sup> First, it fails to account for the studies showing greater belief and knowledge attribution in cases of morally good or morally neutral norm-deviation.<sup>92</sup> Second, the ESEE appears less susceptible to the sorts of otherwise-irrelevant bad character information that, when it increased intentionality- and causality-ascription rates in SEE harm scenarios, led some scholars to conclude that the asymmetries were largely attributable to motivated reasoning.<sup>93</sup>

In short, the relative counter-normativity of P-relevant actions and their outcomes, rather than, or in addition to, the blameworthiness of the agent,

---

88. See Kobick & Knobe, *supra* note 2, at 413; Solan, *supra* note 2, at 525; Holton, *supra* note 62, at 2-4.

89. As explained further in Part IV, *infra*, I am dubious of such a quick move from the descriptive to the normative.

90. This would be an interesting and important observation, and one that fits well within recent accounts of desire-state and causality ascription. See, e.g., Nadler & McDonnell, *supra* note 2, at 301.

91. See generally James R. Beebe, *Do Bad People Know More? Interactions Between Attributions of Knowledge and Blame*, SPRINGER LINK (Sept. 4, 2015), <http://link.springer.com/article/10.1007%2Fs11229-015-0872-4> (published in the journal *Synthese*) (reporting “the results of seven new empirical studies that raise significant challenges to blame-based explanations of the ESEE”).

92. See Alfano et al., *supra* note 69, at 264, 274-76, 281-83 (surveying previous studies and reporting results from two new studies).

93. See Beebe & Jensen, *supra* note 81, at 702-03; Buckwalter, *Gettier Made ESEE*, *supra* note 82, at 368 (reporting results of an experiment indicating that the ESEE “is unlikely to be mediated by a simple desire to blame”).

plays a large role in determining whether an agent is deemed to have known or believed that P. In the following Section, I explain why this is so, not by relying on a philosophical or legal analysis of the concepts in question, but by couching the ESEE within a broader, empirically based account of laypeople's belief-ascription practices and their functional concepts of knowledge and belief.

*B. Bringing the Data Together: A Descriptive Account of Belief State Ascription*

This Section provides an account of what factors influence laypeople's belief-state ascription, using legal examples along the way.<sup>94</sup> It does so in two steps, the first addressing belief ascription (along with awareness and similar concepts), and the second addressing knowledge ascription.<sup>95</sup> These two steps can be summarized as follows: (1) counter-normativity spurs belief formation, and (2) "Actionability" makes (true) belief knowledge.

*1. Ascribing Belief*

Under what circumstances are laypeople likely to ascribe belief that P? In some cases, the answer is easy; perhaps the defendant admits to having believed that P, or his actions would make no sense unless he believed that P. But where the evidence doesn't obviously point one way or another, lay-ascription will be influenced by both (a) how laypeople reason about other people's thought processes, and (b) how laypeople interpret terms like "belief" and "awareness."<sup>96</sup> I address each in turn.

*a) The Norm-Violation / Belief-Ascription Heuristic*

As we have seen, laypeople more often ascribe belief that P to agents whose P-relevant actions were counter-normative. Why would laypeople treat normative valence as relevant to the question of whether or not an

---

94. Although this Section uses legal examples by way of illustration, Part IV, *infra*, explains more fully how these belief state ascription practices relate to criminal law.

95. This second step is where the ESEE is explained. *See infra* Section III.B.2.

96. This Article's account of lay-ascription does not concern trials in which evidence obviously establishes the defendant's mental state, leaving no room for reasonable disagreement. Instead, this Article addresses what might be called "close cases," in much the same way that theories of adjudication often focus on "hard cases" to determine what factors influence judges' interpretations of ambiguous statutory language. Likewise, this Article does not attempt to delineate every single factor that might influence a given juror's belief state ascription even in close cases—e.g., whether the defendant appeared honest while testifying, what the juror ate for breakfast, etc. Instead, it focuses on several underappreciated factors highlighted by recent empirical literature discussed above.

agent believed that P? The answer is that counter-normativity comes with the risk of sanctions, both formal and informal, and that risk makes it comparatively more useful to form true beliefs to the effect that one is violating a norm than to form true beliefs to the effect that one is adhering to a norm.<sup>97</sup> So unless evidence strongly implies otherwise, laypeople presume that others pause to form an accurate belief one way or another before deviating from a norm.<sup>98</sup> In short, laypeople employ what may be called the “norm-violation/belief-ascription heuristic”: all else being equal, insofar as an agent’s action would make it the case that X, where X violates a norm salient to the agent, attribute to the agent the belief that the action at issue would make it the case that X.<sup>99</sup> Where the norm violation is less serious or less salient to the actor, people are less likely to ascribe such beliefs. Conversely, where the norm violation is more serious or more salient, people are more likely to ascribe such beliefs.

*b) Different Kinds of Belief: Occurrent and Dispositional, Thick and Thin*

Laypeople’s ascriptions of belief are responsive not only to how laypeople reason about other minds but also to what they understand terms like “belief” to mean. As it turns out, laypeople understand the concept of “belief” to signify several distinct types of mental states. Experimental epistemologists noted this phenomenon when investigating why—contrary

---

97. Alfano et al., *supra* note 69, at 268-70. Of course, external sanctions need not be the only reason norm-deviations give agents pause. Given agents’ self-interest and moral conscience, it is sensible to assume agents are cautious about breaking norms even where they believe the chance of detection is zero.

98. Recall that the chairman in the original ESEE example stated that he did not care about the environment. Still, it is reasonable for study participants to posit that he cared enough about the *consequences* (to others or, perhaps more likely to himself) of harming the environment to form a true belief about it—hence the greater belief attribution to the chairman in the harm scenario (where his actions are norm deviant) than in the help scenario (where his actions are norm-compliant). *Cf.* Holton, *supra* note 62, at 1.

99. See Alfano et al., *supra* note 69, at 268 (proposing a slightly different formulation of what they label the “Norm-violation/Belief-attribution heuristic”). Generally speaking, the greater the norm violation, the more useful the formation of the true belief in question, and hence the more likely the agent is to form such a true belief. As a result, laypeople will be especially likely to ascribe to an agent belief that P where P represents a more serious or consequential norm violation. Of course, whereas my explanation for the heuristic explains it in terms of how people reason about other minds, a “blame first” error theory might instead posit that the more serious the norm-deviation, the more motivated one is to blame the agent. Either way, the resulting pattern of belief ascription is the same.

to virtually all previous philosophical accounts of knowledge<sup>100</sup> (and contrary to all legal accounts of knowledge)<sup>101</sup>—laypeople sometimes ascribe to an agent *knowledge* that P while simultaneously ascribing to that agent *no belief* that P.<sup>102</sup> How, experimental philosophers wondered, could people ascribe knowledge that P without belief that P? Doesn't knowledge entail belief?<sup>103</sup> The answer turns out to be that knowledge *does* require “belief,” but it requires a more minimal type of belief than laypeople sometimes interpret words like “belief” to denote. This raises an important point: when we ask laypeople to ascribe belief (along with related terms like “awareness” and “conscious disregard”), context will determine which type of belief they take to be at issue.

First, beliefs can be occurrent or dispositional.<sup>104</sup> Occurrent belief is conscious endorsement at a given moment.<sup>105</sup> Dispositional belief is information stored in the mind available for endorsement under typical conditions.<sup>106</sup> Here is one example where laypeople's implicit understanding of belief as either dispositional or occurrent determines their ascriptions of belief and knowledge: Kate panics during her history exam and, despite having repeatedly memorized the date of Queen Elizabeth's death, feels sure she has forgotten.<sup>107</sup> She answers “1603,” thinking it just a

---

100. See Buckwalter & Turri, *supra* note 18, at 8.

101. Although a few philosophers have questioned the belief requirement, *see, e.g.*, Colin Radford, *Knowledge—By Examples*, 27 ANALYSIS 1, 4-5, 11 (1966), legal scholars and judges have not, *see, e.g.*, sources cited *supra* notes 7-8.

102. See Beebe, *A Knobe Effect*, *supra* note 82, at 239; Blake Myers-Schulz & Eric Schwitzgebel, *Knowing That P Without Believing That P*, 47 NOUS 371, 371, 378-80 (2013); Wesley Buckwalter et al., *Belief Through Thick and Thin*, NOUS 1, 6, 18 (Sept. 15, 2013), <http://onlinelibrary.wiley.com/doi/10.1111/nous.12048/pdf>; Buckwalter & Turri, *supra* note 18, at 9 (calling these results “shocking”).

103. See Jonathan Jenkins Ichikawa & Matthias Steup, *The Analysis of Knowledge*, in STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2014), <http://plato.stanford.edu/entries/knowledge-analysis/> (“The belief condition . . . is certainly accepted by orthodoxy.”); Luban, *supra* note 7, at 961 (“[K]nowledge does require belief—I can hardly be said to know something if I don't even believe it”).

104. See Buckwalter & Turri, *supra* note 18, at 9; *see also* Eric Schwitzgebel, *Acting Contrary to Our Professed Beliefs or the Gulf Between Occurrent Judgment and Dispositional Belief*, 91 PACIFIC PHIL. Q. 531 (2010).

105. David Rose & Jonathan Schaffer, *Knowledge Entails Dispositional Belief*, 166 PHIL. STUD. 19, 22 (2013).

106. *Id.* This is only a rough definition of dispositional belief, as the notion of “typical conditions” is ambiguous and may not best describe examples of “thick” dispositional belief. *See* discussion *infra* text accompanying notes 115-116.

107. Myers-Schulz & Schwitzgebel, *supra* note 102, at 375.

guess, but in fact that is the correct answer.<sup>108</sup> Most study participants presented with this vignette assumed “belief” meant *occurrent* belief, and consequently ascribed to Kate knowledge that Queen Elizabeth died in 1603 but not belief that she died in 1603.<sup>109</sup> In follow-up studies, however, participants were primed to consider *dispositional* belief, and consequently ascribed both knowledge and belief.<sup>110</sup> Other vignettes produce similar findings.<sup>111</sup>

To take a legal example, imagine a ski instructor, Hall, who skis down a hill in an unusually dangerous manner, contrary to any advice he would give even an experienced skier. Hall accidentally crashes into another skier, killing the other skier.<sup>112</sup> Did Hall believe his action posed such risk of death (that is, was he aware of the risk so as to be knowing or reckless)? Or did Hall instead have no such belief (that is, did he fail to consider whether, or was he unaware that, his action posed such a risk, so as to be negligent or legally blameless)? If one employs a concept of *occurrent* belief, one is much more likely to conclude that Hall had no such belief; he was likely not consciously considering the risk as he zoomed down the hill. Indeed, because knowledge does not entail *occurrent* belief (i.e., knowledge can involve merely *dispositional* belief),<sup>113</sup> someone primed to consider *occurrent* belief might even say Hall *knew* that skiing in such a manner posed such an unjustifiable risk, but he did not *believe* or was not *aware* of it at the relevant time, much as Kate knew but did not believe that Queen Elizabeth died in 1603.<sup>114</sup> But if one instead employs a concept of *dispositional* belief, one is likely to conclude Hall *did* believe that skiing in that manner posed a grave risk; as a ski instructor, he was likely aware of the risk as a general matter even if he wasn’t consciously considering it at the time.<sup>115</sup> In short, whether the defendant was aware of the risk depends

---

108. *Id.*

109. *Id.*

110. Rose & Schaffer, *supra* note 105, at 34-35.

111. *Id.* at 35-40.

112. This example is based on *People v. Hall*, 999 P.2d 207 (Colo. 2000) (en banc).

113. Rose & Schaffer, *supra* note 105, at 23.

114. *See supra* notes 107-09 and accompanying text.

115. In the *Hall* case, after remand from the Colorado Supreme Court, the trial court’s MPC-based instructions defined recklessness simply as “conscious[] disregard[]” of “a substantial and unjustifiable risk.” Jury Instructions at 14, *People v. Hall*, 59 P.3d 298 (Colo. App. 2002) (No. 97 CR 167). The instructions then stated that “[w]hether a person consciously disregards such a risk may be inferred from either the actors [sic] subjective knowledge of the risk or what a reasonable person with the actor’s subjective knowledge of the risk or what a reasonable person with the actor’s knowledge and experience would have

on whether awareness is occurrent or dispositional, and some factual scenarios and instructions prime different answers to that question.

Second, and independent of the occurrent versus dispositional distinction, beliefs can be either “thin” or “thick.”<sup>116</sup> Thin belief requires merely a “bare cognitive pro-attitude,”<sup>117</sup> the sort of belief people have whenever they think the proposition in question is more likely true than not true. Thick belief requires some sort of emotion or conation,<sup>118</sup> which can be roughly described as an “inner conviction” as to the truth of the proposition in question.<sup>119</sup>

To see the difference between thin and thick belief, consider the following example. George the Geocentrist answers on a test that the earth revolves around the sun. In some sense George acknowledges this on an intellectual level, but, due to deeply felt religious conviction, he believes that the sun revolves around the earth.<sup>120</sup> As it turns out, a significant number of laypeople attribute to George *knowledge* that the earth revolves around the sun but not *belief* that it does (once again, an ascription pattern previously deemed incoherent).<sup>121</sup> They do this because, contrary to legal conceptions of belief as an intellectual probabilistic judgment, the factual scenario implicitly primes them to consider belief in its thick form, which requires some deeper form of mental assent or inner conviction. Other vignettes (not involving religious faith) produce similar findings.<sup>122</sup>

To take a legal example, imagine a case in which a chairman like the one in Knobe’s original vignette is criminally prosecuted for harming the

---

been aware of in the particular situation.” *Id.* at 15. The jury ultimately convicted Hall of negligence, which, in line with the MPC, the jury instructions defined as “fail[ure] to perceive” a substantial and unjustifiable risk. *Id.* at 18; *see* *People v. Hall*, 59 P.3d at 299.

116. Buckwalter et al., *supra* note 102, at 2; *see also* Buckwalter & Turri, *supra* note 18, at 7-14 (surveying experimental findings and concluding that “in light of these results, any program of philosophical or psychological research on ‘belief’ should take into account the difference between thick belief and thin belief”); Dylan Murray et al., *God Knows (But Does God Believe?)*, 166 PHIL. STUD. 83 (2013). Thick beliefs and thin beliefs can be either occurrent or dispositional.

117. Buckwalter et al., *supra* note 102, at 2.

118. *Id.*

119. Murray et al., *supra* note 116, at 102-05.

120. Buckwalter et al., *supra* note 102, at 12-14.

121. *Id.* at 12-14.

122. *See id.* at 14-20; Myers-Schulz & Schwitzgebel, *supra* note 102, at 374-78 (reporting studies resulting in greater ascription of knowledge than belief to, for example, a husband considering whether his wife is cheating on him, and to a subconsciously prejudiced professor considering whether student athletes are just as academically capable as non-athletes).



environment.<sup>123</sup> The question of belief state arises: Did the chairman believe or know the environment would be harmed (recklessness or knowledge), or did he fail to consider whether it would be harmed (negligence or legal blamelessness)? If one employs a concept of thick belief, one is more likely to conclude that the chairman did not believe the policy would harm the environment. Perhaps the busy, flippant chairman didn't care enough about environmental harm to have formed the sort of deeper coming-to-grips-with or acceptance that would satisfy a thick belief requirement. Indeed, employing a concept of thick belief, one might even say the chairman knew the environment would be harmed but did not really believe it. Study participants in the chairman studies, after all, reported greater confidence in ascribing knowledge to the chairman than in ascribing belief.<sup>124</sup> But if instead one employs a concept of thin belief, one is more likely to conclude that the chairman *did* believe that the environment would be harmed—i.e., that on a purely intellectual level, he acknowledged that the environment would likely be harmed.

The distinction between thick and thin belief, though only briefly sketched here, may make a dispositive difference in jurors' mens rea findings in numerous legal contexts, including, for example, cases in which: (1) the defendant purports to have a deep religious conviction relevant to P;<sup>125</sup> (2) the defendant purports to have a cognitive defect related to practical or moral reasoning, such as an inability to "appreciate" the wrongfulness of her actions;<sup>126</sup> (3) the defendant is a group entity (e.g., a corporation), arguably incapable of possessing thick beliefs, or at least less likely than an individual to possess them;<sup>127</sup> (4) the defendant seemingly

---

123. See *supra* note 59 and accompanying text.

124. Beebe, *A Knobe Effect*, *supra* note 82, at 253.

125. See *supra* notes 117-19 and accompanying text; ALEXANDER & FERZAN, *supra* note 14, at 45-46 (discussing cases, cited in LEO KATZ, *BAD ACTS AND GUILTY MINDS: CONUNDRUMS OF THE CRIMINAL LAW* (1987), in which an actor kills someone and claims sincere belief that the victim was a witch, not a human, and legal liability turns in part on whether the actor "knowingly" killed a human); MPC § 3.02(1) (affirmative defense where actor "believe[d]" conduct was necessary to avoid greater harm).

126. See MPC § 4.01 (insanity defense available where defendant cannot "appreciate the criminality [wrongfulness] of his conduct"); M'Naghten's Case, 8 Eng. Rep. 718, 722 (H.L. 1843) (insanity defense available where defendant did not "know the nature and quality of the act he was doing; or, if he did know it, that he did not know he was doing what was wrong").

127. Group agents may be thought incapable of possessing mental states, let alone thick beliefs. But there is mounting evidence that laypeople are comfortable ascribing to them a large array of mental states, some of which may be thick. See, e.g., Avital Mentovich et al., *The Psychology of Corporate Rights* 4-5, 11-12 (Univ. of Chi. L. Sch. Pub. L. & Legal Theory

engaged in “wishful thinking” or some similar form of self-deception;<sup>128</sup> or (5) the defendant’s actions appear to have arisen from sub-conscious prejudice.<sup>129</sup>

To summarize the two main points concerning belief ascription: (a) laypeople tend to ascribe belief that P to an agent taking a P-relevant action where P is norm-deviant,<sup>130</sup> and (b) laypeople’s belief ascriptions are sensitive to what type of belief (occurrent or dispositional, thick or thin) they are primed to consider.

## 2. Ascribing Knowledge

When is a belief that P deemed knowledge that P? Laypeople’s knowledge ascriptions typically reflect the following conception of knowledge<sup>131</sup>: *An agent knows that P if (1) P is true, (2) the agent believes that P (in any sense—dispositionally or occurrently, thickly or thinly), and (3) the agent’s belief that P is “Actionable.”* The first requirement—the

Working Paper Grp., Paper No. 497, 2014), <http://ssrn.com/abstract=2467372> (reporting empirical evidence of laypeople’s ascribing religious convictions to corporations); Adrianna C. Jenkins et al., *The Neural Bases of Directed and Spontaneous Mental State Attributions to Group Agents*, 9 PLOS ONE e105341 (2014), <http://www.plosone.org/article/doi/10.1371/journal.pone.0105341&representation=PDF> (reporting results of fMRI testing revealing that laypeople exhibit similar mental patterns when considering a corporate entity’s mental state and an individual’s mental state). In the corporate context, the difference between thick and thin belief may influence not only the guilt versus innocence question posed to jurors in criminal trials but also the triggering of punitive damages in civil cases. Jurors might interpret a corporation’s having undertaken explicit cost-benefit analyses, for example, as evidence that a corporation engaged in the sort of reflective coming-to-grips-with or premeditation sometimes associated with thick belief. Cf. W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, in CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE 112 (2002).

128. See Myers-Shultz & Schwitzgebel, *supra* note 102, at 378-80 (reporting results from studies concerning wishful thinking); David Sackris & James R. Beebe, *Is Justification Necessary for Knowledge?*, in ADVANCES IN EXPERIMENTAL EPISTEMOLOGY, *supra* note 58, at 190 (same).

129. See Myers-Shultz & Schwitzgebel, *supra* note 102, at 375, 378-80 (reporting results from studies concerning subconscious prejudice).

130. That is, they tend to ascribe *at least thin*, dispositional belief.

131. There are no doubt counter-examples and exceptions to the construal of knowledge offered here. But at the very least, as spelled out in more detail below, it comes closer to capturing laypeople’s concept of knowledge (as that concept operates in both civil and criminal cases, though the focus here is on criminal cases) than does the prevailing account offered by legal theorists and philosophers.

truth condition—is familiar and relatively uncontroversial.<sup>132</sup> The second requirement—the belief condition—is part of the traditional definition of knowledge, though it contains some refinement in light of the different kinds of belief described above.<sup>133</sup> The third requirement—the Actionability condition—is likely new to most readers and requires unpacking.<sup>134</sup>

An agent’s belief is Actionable when (a) it is *relevant* to some salient practical decision, and (b) for purposes of making the practical decision, the agent *ought* (rationally) *to treat its truth as settled*, given the agent’s values, desires, degree of subjective certainty, and practical circumstances. In other words, an agent’s belief that P is Actionable when, in deciding how to act, the agent ought to assume that P.

An agent’s degree of subjective certainty is an important determinant of whether her belief was Actionable. *But*—and this will be important when, in Part IV below, we return to the supposed differences between knowledge and recklessness—the amount of subjective certainty required to make a belief Actionable *depends on context*. More specifically, the requisite amount of subjective certainty varies according to (a) pragmatic features of the agent’s situation, and (b) the normative valence of the course of action at issue. I address each in turn below.

#### *a) Pragmatic Features of the Agent’s Situation*

Whether a true belief was Actionable, and thus counted as knowledge, often depends on the practical options available to the agent at the time she engaged in the P-relevant action in question. Agents often act in situations that don’t require a stark choice between P and Not-P. They can hedge their

---

132. Wesley Buckwalter, *Factive Verbs and Protagonist Projection*, 11 *EPISTEME* 391, 391 (2014).

133. See *supra* notes 104-29 and accompanying text.

134. See Turri & Buckwalter, *supra* note 87, at 4. Two traditional conditions of knowledge will strike philosophers as conspicuously absent: (1) justification, and (2) a proper causal relation between the justificatory evidence and the belief formed. See Ichikawa & Steup, *supra* note 103. Both can be set aside for present purposes. First, justification is largely captured by the Actionability requirement and in any event appears to play a relatively minor role in lay-ascription. See Buckwalter, *Gettier Made ESEE*, *supra* note 82, at 368; Sackris & Beebe, *supra* note 128, at 189 (reporting empirical evidence that laypeople ascribe knowledge to agents whose true beliefs are not “justified” according to the traditional philosophical account of justification); Christina Starmans & Ori Friedman, *The Folk Conception of Knowledge*, 124 *COGNITION* 272, 280 (2012). Second, the “causal relation” condition is relevant to lay-ascription only in an extremely small subset of cases and is rarely implicated in criminal cases. See John Turri et al., *Knowledge and Luck*, *PSYCHONOMIC BULL. & REV.* 378, 382 (2015).

bets because they have courses of action available that allow for a probability assessment to influence the action in question. When such alternative courses of action are available, an agent's bare belief that P (as opposed to her more nuanced probability assessment) is less likely to be Actionable and is thus less likely to be deemed knowledge that P.<sup>135</sup>

To take a legal example, imagine a drug "mule" who was handed a briefcase and told to transport it for money.<sup>136</sup> She believed it 60% likely that the briefcase contained drugs. She wouldn't have transported the briefcase if she had been 100% certain that it contained drugs, but given her view of the odds and the money involved, she decided to transport the briefcase, which, as it turns out, did in fact have drugs inside. She believed there were drugs inside, but did she *know* there were drugs inside?

Imagine two scenarios. In Scenario 1, the briefcase was locked and she had no means of opening it. She had no way of obtaining greater certainty, so in deciding what to do, she should have proceeded on the assumption that there were drugs inside. In other words, her belief that there were drugs inside was Actionable. Therefore, regardless which course of action she ultimately decided on—transport it or don't transport it—she will likely be deemed to have known that there were drugs inside.<sup>137</sup> In effect, laypeople treat the question, "Did she know there were drugs in the briefcase?" as meaning, "*For practical purposes, did she know there were drugs in the briefcase?*"<sup>138</sup>

In Scenario 2, the briefcase is not locked and she is aware of an available low-cost alternative course of action: she could look inside the briefcase. In

---

135. See generally Turri & Buckwalter, *supra* note 87.

136. Cf. *infra* notes 180-98 and accompanying text (discussing the relationship between knowledge and willful ignorance).

137. I use the word "likely" to once again highlight that the factor affecting knowledge ascription here influences, but does not fully dictate, an ascription of knowledge. As noted previously, there are plenty of additional factors that would bear on whether the drug mule would be deemed to have knowledge. As with the study of judicial behavior—or any human behavior, for that matter—this Article does not attempt to account for every such possible influence. Cf. *supra* note 96. Additionally, the phrase "will be deemed to have knowledge," highlights that the important thing here, as elsewhere in this Article, is whether jurors would call a given mental state "knowledge" in every-day life or in court, not whether the mental state lives up to some other more absolute concept of knowledge that diverges from lay usage. See *supra* note 58 and accompanying text; *infra* note 211 and accompanying text.

138. For empirical literature concerning the related issue of stakes to the agent of being wrong about P, and how those stakes impact knowledge ascription, see Pinillos & Simpson, *supra* note 87; Chandra Sekhar Sripada & Jason Stanley, *Empirical Tests of Interest-Relative Invariantism*, 9 EPISTEME 3 (2012); Turri & Buckwalter, *supra* note 87, at 25 (concluding that the stakes effect on knowledge ascription is fully mediated by Actionability).

Scenario 2, given her preferences and the courses of action available to her, she should not treat the truth of the proposition, “there are drugs inside,” as settled. Instead, she should act as if “there is a 60% chance that there are drugs inside”—i.e., she should look inside the briefcase. Whereas in Scenario 1 there was no practical difference between acting as if there were drugs inside and acting as if there was a 60% chance there were drugs inside, in Scenario 2 there is a difference. Because she has this alternative action available in Scenario 2 but not in Scenario 1, her bare belief that there were drugs in the briefcase was Actionable in Scenario 1 but not in Scenario 2, and is thus, all else being equal, more likely to be deemed knowledge in Scenario 1 than in Scenario 2.

As the drug mule example illustrates, one important alternative course of action that affects Actionability is the possibility of considering further whether P or Not-P, either through consulting additional evidence or through engaging in additional deliberation. Sometimes, though, that alternative course of action is unavailable due to time constraints. An agent’s time constraints thus sometimes affect Actionability and thereby affect knowledge ascription.

For example, in one experiment participants were asked about a medical student who, due to an extreme shortage in hospital personnel, is put in charge of a patient and must decide which of three medications to administer.<sup>139</sup> In one scenario, she has months to decide, and she picks the right one. In another she has two minutes to decide, and she picks the right one. In both, she believes the medication she picks is the right one. Participants were significantly more likely to say the student “knew” the medication she picked was the right one when she had two minutes to deliberate rather than several months.<sup>140</sup> Another experiment found the same effect when asking about whether a college student who cares a lot about his paper “knows” his paper contains no misspellings, where in one scenario he has five minutes to proofread it and in the other he has two weeks.<sup>141</sup> In short, at least in some circumstances, all else being equal, the less time you have to consider whether P, the more likely your belief that P is to be deemed knowledge that P.

More generally, where for practical purposes one must act as if P or Not-P, one’s belief that P or that Not-P is deemed knowledge (so long as it turns out to be true), even if one had lingering subjective uncertainty. In other

---

139. Joseph Shin, *Time Constraints and Pragmatic Encroachment on Knowledge*, 11 *EPISTEME* 157, 160-61 (2014).

140. *Id.* at 163.

141. *Id.* at 165.

words, under sufficient practical constraints, knowledge ascription tends to collapse into (true) belief ascription because mere belief, under sufficient practical constraints, is Actionable.

That said, such practical constraints are often not present; agents often have additional time to deliberate, additional evidence to consult, etc. When the agent is not so practically constrained, the amount of subjective certainty necessary to make a belief Actionable often depends on the normative valence of the salient action to which that belief is relevant. Let's now briefly turn to those normative considerations.

*b) Normative Valence of the Course of Action at Issue*

Where moral norms are salient, they predominate in the Actionability analysis. If, and to the extent that, a P-relevant action would deviate from a moral norm in the event P is true, the agent's belief that P is likely to be deemed Actionable.<sup>142</sup> In other words, the worse (in terms of morality) that P being true would make an agent's action, the more likely that the agent's belief that P constitutes knowledge that P. That's because, where P is moral norm-deviant, one need not have very great subjective certainty that P in order to make it the case that one ought to act as if P—i.e., in order to make it the case that one's belief that P is Actionable.

Consider again the chairman example.<sup>143</sup> Imagine that the chairmen in the harm and help conditions each believe the vice president's testimony, thinking it 60% likely that the new program will help or harm the environment as the case may be. Why might the belief be considered Actionable, and hence knowledge, in the harm condition but not the help condition? Because the stakes, in terms of potential moral norm-deviation, are higher in the harm condition than in the help condition.<sup>144</sup> The greater the moral norm-deviation, the lower the subjective certainty required to make the belief knowledge. This same pattern accounts for the numerous experimental findings discussed and cited above, which go far beyond the chairman example.<sup>145</sup> A similar explanation even makes sense of the

---

142. See Beebe & Buckwalter, *supra* note 82, at 494; Buckwalter, *Gettier Made ESEE*, *supra* note 82, at 380.

143. See *supra* note 59 and accompanying text.

144. In the help condition, whether or not the chairman's belief turns out to be correct and the environment is helped, the outcome will not violate a moral norm, since merely failing to actively help the environment is not moral norm-deviant. See, e.g., Holton, *supra* note 62, at 3-4; Solan, *supra* note 2, at 524-25. But in the harm condition, the moral stakes are higher, since approving a policy that harms the environment is moral norm-deviant.

145. See *supra* Section III.A.

otherwise-puzzling ESEE findings in cases of morally positive and morally neutral norm-deviation,<sup>146</sup> though those cases are less directly relevant to criminal law.<sup>147</sup>

*IV. Surprising Divergences Between Mens Rea Theory and Practice:  
What They Are, Why They Matter, and How to Begin Addressing Them*

A brief recap is in order. Part II examined possible ways to construe the belief states invoked in criminal jury instructions. Those instructions contain numerous ambiguities concerning the make-up of the belief states they tell jurors to ascribe. Commentators offering descriptive and normative accounts of criminal mens rea concepts have sought to explain and resolve some of these ambiguities, though without recourse to empirical findings concerning lay-usage.

---

146. See *supra* notes 69-70 and accompanying text. Where there is no moral duty at issue, and P's being true would make an action go against one's self-interest or would make an action go against convention, for example, one's belief that P requires less certainty in order to be Actionable than it would require if it were norm-compliant. See *supra* notes 69-70 and accompanying text.

Consider, as an example of the influence of non-moral norms, a long-standing puzzle in the philosophical literature concerning lotteries: If a lottery ticket holder is aware that there is only a one-in-a-million chance her ticket is a winner, why are laypeople hesitant to ascribe to her, prior to the announcement of the winning number, knowledge that her ticket is a loser? See JOHN HAWTHORNE, *KNOWLEDGE AND LOTTERIES* (2004); Enoch & Fisher, *supra* note 57, at 573-74 (using the lottery example as an allegedly "nonpractical" illustration of laypeople's aversion to ascribing knowledge based on naked statistical evidence); John Turri & Ori Friedman, *Winners and Losers in the Folk Epistemology of Lotteries*, in *ADVANCES IN EXPERIMENTAL EPISTEMOLOGY*, *supra* note 58, at 45. My analysis above suggests an answer: the lottery ticket holder, even if she believes her ticket is almost certainly a loser, should not act on the assumption that it is a loser, but instead should act on her more nuanced probability assessment—i.e., that she has a ticket with a one-in-a-million chance of being a winner. There is a real, salient practical difference between these two ways of acting: in the latter, but not the former, she should hold onto the ticket and check, when the winning number is announced, to see whether her ticket is a winner. That is the rational course of action for a lottery player to take, given her preferences (she has, after all, bothered to obtain a ticket!). Because she should not ignore or discard the ticket prior to the announcement of the winning number, her bare belief that the ticket is a loser is not Actionable, and she is thus unlikely to be ascribed knowledge that the ticket is a loser despite the overwhelming odds and her awareness of those odds.

147. Criminal law typically concerns deviations from salient *moral* norms. Nonetheless, morally neutral and morally positive norm-deviation cases are worth noting because of their implications (a) for error theories and competence theories (i.e., whether the SEE, ESEE, and other findings are wholly attributable to blame-based motivated reasoning), and (b) for areas of criminal law that do not mirror moral norms (e.g., *malum prohibitum* offenses).

Part III examined the ways laypeople construe and ascribe belief, knowledge, and other belief states. Broadly speaking, it demonstrated two things. First, belief-state ascription (including knowledge ascription) is much more dependent upon normative and pragmatic features of the agent's situation than legal and philosophical accounts have traditionally assumed.<sup>148</sup> Second, beliefs vary in kind—they can be occurrent or dispositional, thick or thin—and laypeople's belief-state ascription often depends on which type of belief they assume they're being asked about.

This Part returns to the specific ambiguities in mens rea concepts described in Part II and examines, in light of the lay-ascription practices outlined in Part III, how jurors are likely to resolve them. That is, it examines how jurors are likely to construe and ascribe belief, knowledge, and other legally relevant belief states in criminal trials. It focuses on several areas of systematic divergence between, on the one hand, what scholars, judges, and code drafters *think* mens rea terms mean and, on the other hand, what those terms *actually* mean as implemented by jurors adhering to jury instructions. After describing each divergence, this Section explains why the divergence matters and how it might be addressed. Finally, Section IV.E argues that current institutional mechanisms are ill-designed to note and remedy problematic mens rea instructions, highlighting the need for further research in this area.

#### A. *The Vanishing Distinction Between Knowledge and Recklessness*

Recall the scholarly consensus concerning the difference between knowledge and recklessness under the MPC. Knowledge that P is thought to be an “invariant” and “purely descriptive” mental state: it requires (a) true belief that P, plus (b) some contextually invariant level of certainty that significantly exceeds a 50% probability assessment, (c) without regard to any other factors.<sup>149</sup> Even outside the MPC, this closely tracks the conventional wisdom regarding what “knowledge” means when used in

---

148. This conclusion coheres with a more general discovery central to much of experimental philosophy, namely, the pervasive impact of evaluative and pragmatic factors on laypeople's ascription of concepts previously thought to be less context-dependent and evaluative and more purely descriptive. See Knobe, *supra* note 63 (reviewing the literature).

149. See Simons, *Should the Model*, *supra* note 7, at 189; Ferzan, *supra* note 24, at 2529-30 (distinguishing “mechanical” mens rea concepts such as purpose, knowledge, and willful blindness under the MPC, from “evaluative” mens rea concepts such as MPC recklessness and negligence); Simons, *Understanding the Topography*, *supra* note 41, at 248 n.51; Gardner, *supra* note 42, at 725 (MPC “purpose” and “knowledge” are fully descriptive, while “recklessness” has both descriptive and evaluative aspects); Alexander, *supra* note 41, at 940.



ordinary speech and thus when it appears undefined in statutes and jury instructions, as frequently happens in federal criminal cases.<sup>150</sup> Recklessness, on the other hand, is a “variant” and “evaluative” mental state.<sup>151</sup> What constitutes an “unjustifiable” (if not also a “substantial”)<sup>152</sup> risk shifts according to normative and pragmatic aspects of the actor’s situation—i.e., according to whether, “considering the nature and purpose of the actor’s conduct,” taking the risky action represents a gross deviation from reasonable standards of conduct.<sup>153</sup>

This distinction between knowledge as an invariant, descriptive mental state and recklessness as a variant, evaluative mental state falls apart in practice. As discussed in Part III, laypeople construe “knowledge” that P as what one might call “practical certainty” that P, where practical certainty means something like: “for practical purposes, certain enough that P to make it the case that someone in the agent’s situation should act as if P.”<sup>154</sup> In other words, laypeople treat knowledge as a contextually variant and evaluative mental state much like recklessness.<sup>155</sup>

To review, here is how the normative and pragmatic features typically associated with recklessness factor into knowledge ascription. Where circumstantial evidence of mental state is a close call, and P turned out to be true, jurors will employ the norm-violation / belief-ascription heuristic, tending to construe the defendant as having had awareness of a risk that P,<sup>156</sup> and as having had a belief that P.<sup>157</sup> Any type of belief—whether

---

150. See *supra* notes 46-48 and accompanying text.

151. See *supra* notes 41-42, 54 and accompanying text.

152. See *supra* notes 34, 39-41 and accompanying text, *infra* note 199.

153. MPC § 2.02; *cf. id.* cmt. 3, at 237 (“[T]he acceptability of a risk in a given case depends on a great many variables.”).

154. See *supra* Section III.B.2. Interestingly, the MPC defines knowledge as to results in terms of being “aware that it is *practically certain* that [one’s] conduct will cause such a result.” MPC § 2.02(2)(b)(ii) (emphasis added). But there is no indication in the MPC’s drafting history or the criminal theory literature that “*practically certain*,” as used in that definition, refers to Actionability, as opposed to meaning “*almost certain*.”

155. For an overview of evolutionary game theory literature suggesting the usefulness of such a practical concept of “knowledge,” see James Beebe, *Social Functions of Knowledge Attributions*, in KNOWLEDGE ASCRIPTIONS 220 (Jessica Brown & Mikkel Gerken eds., 2012).

156. For this reason, it is unsurprising that, in a recent study, jury-eligible laypeople tasked with matching descriptions of mental states to their proper MPC mens rea terms mistakenly labeled negligence as recklessness 31% of the time. See Francis X. Shen et al., *Sorting Guilty Minds*, 86 N.Y.U. L. REV. 1306, 1343 (2011). They did this despite being provided the MPC definitions of each, and despite receiving direct descriptions of the defendant’s mental state—rather than having to infer it from circumstantial evidence—which described the mental state in similar terms to the MPC definitions with which they

occurrent or dispositional, thick or thin—will suffice for knowledge so long as it was *Actionable*.<sup>158</sup>

The question of whether a belief that P was Actionable can be restated as follows: Would a reasonable person in the defendant's situation have acted as if P?<sup>159</sup> The question can be fleshed out in terms of the pragmatic and normative considerations outlined in Part III. For example, if the defendant was forced to act as if either P or Not-P, and no alternative “hedging” courses of action were available—e.g., if there was no time to deliberate, or no additional evidence to obtain before deciding how to act—then the defendant's bare belief that P was Actionable. The reasonable person in such a situation, believing that P, would act as if P.<sup>160</sup> What if the defendant had other available courses of action? In that case, moral norms will play a decisive role in determining Actionability.<sup>161</sup> The defendant's belief that P was Actionable insofar as (1) the defendant's P-relevant action would be moral norm-deviant in the event that P turned out to be true (i.e., in the event the risk is borne out), and (2) the defendant had no overriding reason (moral, prudential, or otherwise) *not* to simply take P as settled and act as if P.<sup>162</sup> Where P then turns out to be true—i.e., where the risk of which the defendant was aware is borne out—the defendant was both reckless and knowing. Virtually the same considerations determined both attributions.<sup>163</sup>

What role, then, does the defendant's degree of subjective certainty actually play in determining whether he possessed knowledge or instead

were to be matched. *Id.* at 1330-33. In the face of even a small amount of ambiguity concerning whether the defendant was aware of the risk, people can be expected to resolve the ambiguity in favor of a finding of awareness, even where the action at issue represented only a trivial, non-criminal norm-deviation. *See id.* at 1330 (providing the language used to signify negligence in the study's vignettes, such as “carelessly,” “wasn't paying attention,” and “hurriedly”); *supra* Section III.A.1.

157. *See supra* Section III.A.1.

158. *See supra* notes 132-35, 138 and accompanying text.

159. *Cf.* MPC § 2.02(2)(c) (defining recklessness in terms of “the standard of conduct that a law-abiding person would observe in the actor's situation”).

160. *See supra* Section III.B.2.a.

161. *See supra* Section III.B.2.b. They may not play such a role in the case of *malum prohibitum* crimes and in other areas where jurors do not think the criminalized act at issue is immoral. In those cases, prudential and descriptive norms will play a more significant role. *See supra* notes 69-70, 146-47 and accompanying text.

162. *Cf.* MPC § 2.02(2)(c) (including in the definition of recklessness consideration of “the nature and purpose of the actor's conduct”); *supra* note 134 and accompanying text.

163. Indeed, given that recklessness requires a “gross deviation” from reasonable standards of conduct, it may in some instances be a higher bar than knowledge, which has no such requirement. *Compare* MPC § 2.02(2)(c) *with id.* § 2.02(2)(b).

mere belief? Degree of subjective certainty is important only insofar as it bears on Actionability. Any requisite certainty threshold for knowledge, if there is one, depends on context (e.g., normative and pragmatic factors, type of belief primed) and can get at least as low as 50%.<sup>164</sup>

But, one might object, surely jury instructions convey a different “legal definition” that distinguishes knowledge from recklessness and renders jurors’ prior natural language-based understanding irrelevant. Not so. Many federal instructions offer no definition of “knowledge.”<sup>165</sup> And under the MPC, recall the open questions concerning what counts as a “substantial and unjustifiable” risk that P (for recklessness), as opposed to a “high probability” that P or a “practical certainty” that P (for knowledge). The terms “substantial,” “high,” and even “practical,” are heavily context-sensitive, as shown not only by commonsense reflection on ordinary language but also empirical linguistic research.<sup>166</sup> Such vague directives leave jurors to construe the term “knowledge” in court much as they do out of court, which is to say, in a way that reflects the sorts of pragmatic and normative considerations that recklessness tracks. Indeed, they are told to do nothing to the contrary. For this reason it is no surprise that, in a recent study, jury-eligible laypeople tasked with matching descriptions of mental states to their proper MPC mens rea terms, even when provided the MPC’s

---

164. To be clear, I do not mean to imply that laypeople will always or typically think in such numerical terms without prompting, nor that they will typically think in terms that are reducible to numerical probabilities without losing important information potentially bearing on the defendant’s culpability.

Moreover, where thick belief is primed, laypeople may even allow the probability sufficient for a belief that fulfills knowledge’s belief requirement to, in effect, dip below 50%. That is, laypeople might think a thick belief that goes against the evidence—even in the defendant’s own estimation—is nonetheless Actionable, especially where the belief in question is perceived to be a useful or morally praiseworthy belief to have. *See* Sackris & Beebe, *supra* note 128, at 175-87 (reporting experiments in which subjects ascribed knowledge to (1) a father who, despite overwhelming evidence against his daughter, believes she did not commit a crime, and (2) a husband who, despite the evidence against his wife’s prospects for surviving cancer, believes she will survive). *But see* Husak & Callender, *supra* note 8, at 38-39 (reasoning that awareness of a less than 50% chance that P entails “actual belief” that Not-P).

165. *See supra* notes 43-45.

166. *See e.g.*, Paul Egre & Florian Cova, *Moral Asymmetries and the Semantics of “Many”*, 8 SEMANTICS & PRAGMATICS 13 (2015), [http://semprag.org/article/view/sp.8.13/pdf\\_8\\_13](http://semprag.org/article/view/sp.8.13/pdf_8_13); Stephanie Solt, *Vagueness in Quantity: Two Case Studies from a Linguistic Perspective*, in UNDERSTANDING VAGUENESS: LOGICAL, PHILOSOPHICAL, AND LINGUISTIC PERSPECTIVES 157 (Petr Cintula et al. eds., 2011); Shalom Lappin, *An Intensional Parametric Semantics for Vague Quantifiers*, 23 LINGUISTICS & PHIL. 599 (2000).

mens rea definitions, fared only slightly better than chance at distinguishing knowledge from recklessness, and indeed, chose “knowledge” more often than “recklessness” when recklessness was the correct answer.<sup>167</sup> In practice, there is virtually no difference between recklessness and knowledge.

This lack of difference matters for at least three reasons. First, since “knowledge” in criminal law is a more capacious category than the conventional wisdom suggests, laypeople likely “over”-ascribe knowledge to criminal defendants. That is, laypeople ascribe knowledge in more cases than would be warranted on the more lofty, rarefied understanding of knowledge assumed in the scholarly literature, MPC commentaries, and case law, and they do so while adhering to current jury instructions.

Second, a prosecutor’s decision to charge a defendant with a crime requiring knowledge, as opposed to a crime requiring recklessness, can drastically increase the defendant’s punishment without any corresponding difference in the actual underlying mental state the jury will be asked to ascribe. For example, in Colorado, an MPC jurisdiction, a knowing homicide carries a mandatory sentence of between sixteen and forty-eight

---

167. See Shen et al., *supra* note 156, at 1343 (reporting that subjects: (a) correctly labeled instances of knowledge only 50% of the time, inaccurately labeling them as “reckless” 30% of the time; and (b) correctly labeled instances of recklessness only 40% of the time, inaccurately labeling them knowledge 42% of the time); *id.* at 1346 n.94 (noting that “[i]t is not the purpose of this Article to explore the many reasons why subjects might have difficulty at this K/R boundary”). Study participants fared better in distinguishing negligence from recklessness and were much more accurate in sorting other mental states. *Id.* at 1343 (reporting 78% and 88% accuracy rates in matching “purposely” and “blameless” vignettes, respectively, with their proper MPC definition). In a follow-up study, “[r]educing the communicated probability (e.g., from ‘very likely’ to ‘some risk’ and from ‘likelihood’ to ‘real risk’)” in the study vignettes’ descriptions of reckless protagonists’ mental state, “improved the ability of participants to accurately identify the mental state”—i.e., to match it to the correct MPC definition. Matthew R. Ginther et. al., *The Language of Mens Rea*, 67 VAND. L. REV. 1327, 1356 (2014). Still, the authors write,

Even in our best case, only 59% of subjects are accurately identifying R[ecklessness] scenarios. . . . About 70% of these misidentifications are subjects believing that a[] R[ecklessness] scenario demonstrates knowing conduct on the part of the protagonist. We are still left with the basic conclusion we reached in the original study: laypeople have great difficulty identifying and distinguishing reckless and knowing behavior.

*Id.* at 1359.

years,<sup>168</sup> while a reckless homicide carries a non-mandatory sentence of two to six years.<sup>169</sup>

Third, for legislatures drafting criminal codes and judges deciding which mental state to read into statutes that fail to specify a mens rea, the choice between knowledge and recklessness typically makes little difference to the guilt/innocence determination under the current jury-instruction regime. Drawing the line at knowledge, rather than recklessness, is in this respect more arbitrary and inconsequential than has been previously assumed.<sup>170</sup> Unfortunately, while the mental state these terms invoke may not differ in practice, the consequences for defendants at the sentencing stage may be great, presumably owing to legislatures' or courts' failure, when determining appropriate sentences, to appreciate the lack of difference in the underlying mental states.<sup>171</sup>

If the current distinction between recklessness and knowledge is problematic, there are two broad ways criminal law might address the problem: abandon the distinction or salvage it. Abandoning the distinction would be relatively easy to implement. Salvaging the distinction is more difficult, in that it calls for some sort of clarification in jury instructions concerning knowledge. One way would be to explicitly quantify the degree

---

168. COLO. REV. STAT. §§ 18-3-103(1), 18-3-103(3)(a), 18-1.3-406 (2015).

169. *Id.* §§ 18-3-104, 18-1.3-401(1)(a)(V)(A).

170. The Supreme Court recently examined the applicable mens rea for the federal statute criminalizing threats of violence. *See* *Elonis v. United States*, 135 S. Ct. 2001 (2015). The statute at issue was silent as to mens rea. The Court reasoned that negligence was insufficient for conviction under the statute, but the Court did not reach the question of whether recklessness would suffice or whether instead knowledge or purpose was required. *Id.* at 2017. In a partial dissent, Justice Alito criticized the majority's decision not to specify the correct mens rea, noting "regrettable consequences" of the Court's incrementalism. *Id.* at 2014 (Alito, J., dissenting in part and concurring in part) ("If purpose or knowledge is needed and a district court instructs the jury that recklessness suffices, a defendant may be wrongly convicted. On the other hand, if recklessness is enough, . . . a guilty defendant may go free."). Justice Thomas's dissent implies that the majority coalition could not agree on whether recklessness should suffice where a statute is silent. *See id.* at 2028 (Thomas, J., dissenting) ("Given the majority's ostensible concern for protecting innocent actors, one would have expected it to announce a clear rule—any clear rule. Its failure to do so reveals the fractured foundation upon which today's decision rests."). In short, the Court couldn't reach consensus concerning whether recklessness sufficed or whether conviction under the statute instead required knowledge, a debate that is likely to persist in the coming years as lower courts decide which mens rea requirement to impose where federal statutes are silent.

171. *See, e.g.,* *United States v. Smith*, 544 F.3d 781, 786 (7th Cir. 2008) (concluding that the Armed Career Criminal Act's enhanced penalties are triggered by a mens rea of knowledge but not recklessness).

of subjective certainty necessary for a belief to count as knowledge—say, 90% certainty.<sup>172</sup> This would make knowledge a less variant and evaluative mental state. It would also force jurors to consider belief in the probabilistic terms in which legal scholars typically describe it, rather than the thick type of belief that factual circumstances sometimes prime jurors to consider. In both of these respects, knowledge would be more effectively distinguished from recklessness than it currently is.<sup>173</sup>

In the end, one's preference for keeping and clarifying the distinction or instead for eliminating it properly depends on a host of normative considerations beyond merely the difficulty, and current failure, in effectively distinguishing the two concepts. In that vein, a cautionary note is in order concerning how to approach the relevant normative considerations.

Some might be tempted to base their preferred solution on whether they endorse a competence or instead an error theory as an explanation for the findings reported in Part III. At least, the literature concerning the SEE showed such an inclination. Recall that in the SEE literature, competence theorists, who believe that the SEE reveals something about the concept of intentionality, of causality, etc., (rather than about the infiltration of bias and motivated reasoning), seemed apt to recommend that the law more closely track lay-judgments of intentionality and causality.<sup>174</sup> In other words, competence theorists said, “now that we better understand what intentionality is (or what causal responsibility is, or whatever), let's make sure that where the law invokes these concepts, it uses them in accordance with what we've discovered to be their true meaning.”

Error theorists, on the other hand, who think that the SEE is evidence of motivated reasoning or some other cognitive defect, proposed that law ought to seek to eliminate such biases in an effort to maintain the pure descriptiveness of concepts like intentionality and causality.<sup>175</sup> That is, error theorists looked at the data and said, “now that we better understand what's getting in the way of laypeople accurately applying the concept of intentionality (or causal responsibility, or whatever), let's make sure that when the law invokes these concepts, it takes pains to prevent laypeople's

---

172. *C.f.* Simons, *Should the Model*, *supra* note 7, at 183 (making a similar proposal on different grounds).

173. That said, quantifying subjective certainty in this way would have downsides. For example, jurors may have trouble applying probability percentages.

174. *See supra* note 72 and accompanying text.

175. *See supra* notes 77 and accompanying text.

prejudices from getting in the way of their accurately applying these concepts.”

My cautionary note is that this method of reasoning moves too hastily from the descriptive to the normative.<sup>176</sup> It focuses too strictly on conceptual analysis—on making sure that the law tracks a given term, whatever that term turns out to mean—at the expense of the underlying normative considerations that should be driving the analysis.<sup>177</sup> In the end, whether the ESEE, along with the other findings discussed in Part III, reveals something about a given belief state *properly understood*, or instead reveals something about how laypeople *misconstrue or misapply* that belief-state concept, the bottom-line normative issue is whether the outcomes in criminal trials—and thus in some sense the law itself—match up with the proper aims of criminal law.

*B. A Peculiar Difference Between Willful Ignorance and Knowledge*

What little is left of the distinction between recklessness and knowledge has important implications for ongoing debates concerning the relation between “willful ignorance” and “knowledge,” and the propriety of allowing willful ignorance to satisfy statutory knowledge requirements. Willful ignorance has two broad requirements: “(1) the defendant must subjectively believe that there is a high probability that a fact exists and (2) the defendant must take deliberate actions to avoid learning of that fact.”<sup>178</sup> The “deliberate action” may be as minimal as “a cutting off of one’s normal curiosity by an effort of will.”<sup>179</sup> A paradigmatic example of a willfully

---

176. For this reason, an ultimate conclusion as to whether to keep or discard the distinction between knowledge and recklessness is outside the scope of this Article. But if that is so, then why bother discussing the error versus competence theory divide at all? In part, as indicated, simply to head off a problematic means of reasoning through the normative implications of the theories discussed in Part III. But also because, for those who favor maintaining the distinction between knowledge and recklessness, the choice of an error or a competence theory may rightly impact the *means* they propose for clarifying the distinction. By highlighting the merits of a competence theory, this Article seeks to broaden the discussion of solutions—which, in similar debates, sometimes assumes an error theory and then gestures toward difficult “de-biasing” mechanisms—to include a relatively simple, if only partial, solution: altering jury instructions in the hopes that people will follow them.

177. To be sure, lay-usage, along with conceptual analysis that tracks it, can be highly relevant to a host of important issues (e.g., statutory interpretation, fair notice, crafting implementable legal rules, etc.). My point here is simply that such analysis is less relevant to the more fundamental question of what criminal law ought to prohibit in the first place.

178. *Global-Tech Appliances, Inc. v. SEB S.A.*, 131 S. Ct. 2060, 2070 (2011).

179. *United States v. Giovannetti*, 919 F.2d 1223, 1229 (7th Cir. 1990). The MPC, which does not define willful ignorance as separate from knowledge, does not require the

ignorant defendant is the drug “mule” who, despite her curiosity, fails to check the contents of the briefcase she is paid to transport.<sup>180</sup>

Most scholars distinguish willful ignorance from knowledge,<sup>181</sup> often deeming it a species of recklessness.<sup>182</sup> But in doing so, despite purporting to track lay-usage or the legal meaning of the concept of knowledge, they appeal to supposedly intuitive concepts of knowledge that track neither.<sup>183</sup> Husak and Callender, for example, note (quite plausibly) that “[i]n most cases, the (sincere) willfully ignorant defendant would admit that he believes p, but would deny that he knows p.”<sup>184</sup> But they then proceed “on the assumption that knowledge consists of some kind of externally justified true belief,” and reason that “[t]he foremost question in deciding . . . whether the willfully ignorant defendant possesses genuine knowledge, is whether his belief in the incriminating proposition is justified.”<sup>185</sup> In the end, they conclude that willful ignorance is distinct from knowledge: “Many willfully ignorant defendants will lack sufficient justification for p, and thus will not know p.”<sup>186</sup> The problem with this sort of approach is that it stresses “justification”—a necessary condition for knowledge under

---

“deliberate action” prong, and adds a requirement that the defendant not “actually believe[]” that the inculpatory proposition is not true. MPC § 2.02(7). Some non-MPC instructions further require that the defendant’s ignorance was the product of the defendant’s desire to escape legal liability. *See, e.g.,* United States v. Willis, 277 F.3d 1026, 1031-32 (8th Cir. 2002).

180. *See, e.g.,* Husak & Callender, *supra* note 8, at 37; ALEXANDER & FERZAN, *supra* note 14, at 34.

181. *See, e.g.,* Charlow, *supra* note 27, at 1390; Robbins, *supra* note 8, at 226; Frans J. Von Kaenel, *Willful Blindness: A Permissible Substitute for Actual Knowledge Under the Money Laundering Control Act?*, 71 WASH. U. L.Q. 1189, 1212-13 (1993); Husak & Callender, *supra* note 8, at 51; Jessica A. Kozlov-Davis, *A Hybrid Approach to the Use of Deliberate Ignorance in Conspiracy Cases*, 100 MICH. L. REV. 473, 482-83 (2001); *see also* *Global-Tech*, 131 S. Ct. at 2072 (Kennedy, J., dissenting) (“Willful blindness is not knowledge; and judges should not broaden a legislative proscription by analogy.”).

182. *See, e.g.,* ALEXANDER & FERZAN, *supra* note 14, at 34 (“The prototypical willfully blind actor is, of course, reckless.”).

183. *Cf.* Garvey, *supra* note 21, at 370 (“Scholars have offered at least three accounts of willful ignorance. These accounts differ because each begins with a different analysis of the concept of knowledge.”).

184. Husak & Callender, *supra* note 8, at 46; *cf.* MPC § 2.02(7) cmt. 9, at 248 (describing “deliberate ignorance” under the MPC as involving “the case of the actor who is aware of the *probable* existence of a material fact but does not determine whether it exists or does not exist”) (emphasis added).

185. Husak & Callender, *supra* note 8, at 46-47.

186. *Id.* at 51.



traditional philosophical accounts,<sup>187</sup> but one that plays little role in laypeople's knowledge ascriptions and appears nowhere in jury instructions.<sup>188</sup> Much scholarly discussion of willful ignorance—concerning both whether it is the same as knowledge and, if not, whether it is as culpable as knowledge<sup>189</sup>—is similarly premised on conceptions of knowledge that are out of sync with how laypeople understand the term and how jury instructions define it.<sup>190</sup>

Putting those issues aside, it is still true that willful ignorance can be distinguished from knowledge. For example, if the defendant has no belief that P, then he does not know that P and can be reckless or willfully ignorant with respect to P without knowing that P.<sup>191</sup> But if the presence or absence of belief were the only difference between knowledge and willful ignorance, then the debate over the relative culpability of the two mental

---

187. See Ichikawa & Steup, *supra* note 103.

188. See Beebe, *supra* note 155 (providing empirical evidence that, for example, laypeople ascribe knowledge to holders of true beliefs that were formed through hallucination, or that study participants themselves judge as going against the evidence of which the agent was aware). Nor is it clear that the question of whether one's true belief was justified *should* play a role in the relevant culpability determinations. Whereas subjective certainty concerning a risk is typically, and rightly, thought to bear on culpability, epistemic justification bears on it, if at all, only much more indirectly. On the culpability "grading function" of subjective certainty, see Simons, *Statistical Knowledge*, *supra* note 23, at 15-16.

189. See, e.g., Charlow, *supra* note 27, at 1417-18; cf. *Global-Tech Appliances, Inc. v. SEB S.A.*, 131 S. Ct. 2060, 2069 (2011) ("The traditional rationale for this doctrine is that defendants who behave in this manner are just as culpable as those who have actual knowledge.").

190. See, e.g., sources cited *supra* notes 7-8. Likewise, in explaining why willful ignorance is or is not a form of knowledge, the literature often employs arguments that assume normative and pragmatic factors play no proper role in knowledge ascription. See, e.g., ALEXANDER & FERZAN, *supra* note 14, at 34 (stating as a premise that, "If a similar risk imposition would not be deemed 'knowing' if undertaken for good reasons . . . then it is misguided to deem the risk imposition 'knowing' merely because one disapproves of the reasons for undertaking it"); Husak & Callender, *supra* note 8, at 51 (analogizing between two instances of awareness with less than full certainty—only one of which is accompanied by an easy means of obtaining greater certainty—and assuming that the availability of a means of obtaining more certainty is irrelevant to whether the actor's mental state counts as knowledge).

191. Of course, that may be relatively rare, at least in close cases, given laypeople's belief-ascription practices—namely, (1) the norm-deviation / belief-ascription heuristic, and (2) the numerous degrees and types of belief that can support knowledge ascription. Cf. Husak & Callender, *supra* note 7, at 42 ("In most cases, the (sincere) willfully ignorant defendant would admit that he believes p"); MPC § 2.02(7) cmt. 9, at 248 (describing "deliberate ignorance" under the MPC as involving "the case of the actor who is aware of the *probable* existence of a material fact") (emphasis added).

states would be the same as the debate over the relative culpability of knowledge and recklessness. What else might distinguish willful ignorance from knowledge in the typical case, where both mental states involve belief in the inculpatory proposition?<sup>192</sup>

This Article's analysis highlights an additional factor distinguishing some cases of willful blindness from knowledge: the availability of a salient, cheap alternative course of action—e.g., checking inside the briefcase—in the case of willful blindness but not in the case of knowledge. For example, as discussed above in Part III, the drug mule who feels 60% certain there are drugs in the easy-to-open briefcase should not treat as settled her belief that “there are drugs in the briefcase” but rather should act on her more nuanced belief that “there is a 60% chance there are drugs in the briefcase.”<sup>193</sup> There is a real practical difference between the two: the latter belief dictates that she look inside the briefcase, while the former dictates treating the matter as settled and deciding whether to transport drugs.<sup>194</sup> In other words, the ease with which the prototypical willfully ignorant actor could obtain additional information makes that actor's more nuanced probability assessment Actionable while it makes non-Actionable her bare belief that P. Thus, in the rare instances of willful ignorance plus belief that P but not knowledge that P, it is often the availability of an easy means of obtaining certainty—rather than some difference in the defendant's subjective certainty or epistemic justification—that would prevent the defendant's belief that P from being deemed knowledge that P.

This distinguishing factor highlights one type of case in which a willful ignorance instruction would mean the difference between a finding of knowledge and a finding of no knowledge. It may also help focus the inquiry concerning the relative culpability of knowledge and willful ignorance. Ongoing scholarly debate on that topic might fruitfully address what it is about having an available means of obtaining certainty that might make a defendant more culpable at the same time it makes that defendant's belief less likely to constitute “knowledge.”<sup>195</sup>

---

192. See *supra* note 184 and accompanying text.

193. See *supra* Section III.B.2.a.

194. See *supra* Section III.B.2.a.

195. For an account along these lines, see Alex Sarch, *Willful Ignorance, Culpability and the Criminal Law*, 88 ST. JOHN'S L. REV. 1023, 1080-81 (2014) (suggesting that a “duty of reasonable investigation” plays a critical role in those cases in which willful ignorance is as culpable as knowledge).

*C. Implicit, and Outcome-Determinative, Differences in “Belief” and Its Relatives*

Now consider the issue of criminal law’s treatment of “belief,” along with related belief-state concepts such as “awareness,” “conscious disregard,” and “failure to perceive,” which are invoked throughout the MPC, the federal code, and jury instructions in both types of jurisdiction.<sup>196</sup> Recall that jury instructions in MPC and federal jurisdictions typically invoke these terms without defining them, and that commentators construe them as denoting thin, typically dispositional, belief states.<sup>197</sup>

The empirical studies surveyed in Part III indicate that when laypeople are not told which type of belief state is at issue, they sometimes implicitly assume it is thick, as opposed to thin, or occurrent, as opposed to dispositional, based on what a given factual scenario or instruction primes them to consider.<sup>198</sup> When legislatures and courts task juries with ascribing a given belief-state concept—“conscious disregard,” “awareness,” “belief,” etc.—the very same term, as it is used in the very same statutory provision, may be systematically construed as denoting a different type of belief state in different cases due to small differences in the factual circumstances and/or jury instructions. This is important because some case outcomes hinge on whether jurors implicitly construe legally relevant belief states as thick or thin and as dispositional or occurrent.<sup>199</sup>

The difference between a finding of belief, awareness, conscious disregard, etc., and a finding of no such belief state can be the difference between a defendant’s being found reckless and a defendant’s being found negligent (or innocent, where negligence is not criminalized).<sup>200</sup> But it can also be the difference between *knowledge* and negligence (or between knowledge and innocence). Indeed, where a notion of thick belief or occurrent belief is primed, it may actually be easier to show that a defendant knew that P than to show that the defendant believed that P: laypeople may implicitly think the defendant had a thin, but not a thick belief, or had a dispositional, but not occurrent belief, and hence find knowledge (which can be found so long as any type of belief is present)

---

196. See *supra* Part II.

197. See *supra* Part II.

198. See *supra* Section III.B.1.b.

199. For examples, see *supra* Section III.B.1.b.

200. See, e.g., Husak, *supra* note 4, at 207-08 (arguing that “the boundary between recklessness and negligence is unclear” due to failure to differentiate between occurrent and dispositional belief); Kimberly Kessler Ferzan, *Opaque Recklessness*, 91 J. CRIM. L. & CRIMINOLOGY 597 (2001).

while simultaneously finding no belief (as they understand the concept in context).<sup>201</sup> While legislators and judges may assume that by implementing a knowledge requirement, they are invoking a belief state that will be less readily ascribed than “mere” belief, that will not always be the case. Once again, knowledge is in some ways a less stringent and more variant requirement than more traditional legal and philosophical accounts would have it, in part because the law is not clear about what type of belief must underlie it.

If the status quo is problematic because the “wrong” type of belief is sometimes primed,<sup>202</sup> what is to be done? Jury instructions could clarify what type of belief is at issue in a given trial, at least where the “wrong” type is otherwise likely to be primed. The statutes on which those instructions are based could do likewise.<sup>203</sup> Of course, insofar as the recommendation is to clarify jury instructions or statutes, a natural next question is: What is the “right” kind of belief for criminal law to specify? This question once again should not be answered by recourse to an error theory or a competence theory and a focus on some singularly “true” meaning of “belief,” “awareness,” etc. Instead, it turns on complex normative considerations outside this Article’s scope.<sup>204</sup> But by staying

---

201. See *supra* Section III.B.1.b.

202. The status quo may also be problematic on grounds of fair notice, consistency of implementation, vagueness and consequent prosecutorial discretion, and other problems familiar in contexts where criminal law’s prohibitions are ambiguous. Cf. J. Kelly Strader, *(Re)Conceptualizing Insider Trading Law*, 80 Brooklyn L. Rev. 1, 4 (2015) (noting that vague mens rea elements raise “two basic due process concerns: denying potential defendants fair notice and emboldening prosecutors to push the law beyond established boundaries”).

203. That is, legislatures or code drafters might specify what they mean by the various belief state terms they use, at least along the axes of occurrent versus dispositional and thick versus thin. They might do this in the language of the statute itself, or through examples found in official commentaries, similar to the “Illustrations” often found in Restatements of the Law. The use of such illustrations, rather than language in the statutory provision itself, might prove necessary insofar as the legislature intends for the type of belief at issue to be different in different factual circumstances. In any event, clarification by the legislature would help address fair notice and related concerns mentioned *supra* note 202.

204. Of course, that is not to say that there aren’t some readily apparent trade-offs involved, whatever conclusion one ultimately draws. Consider, for example, whether awareness should be construed as occurrent or merely dispositional. A construal of awareness as merely *dispositional* risks undermining the distinction between knowledge and recklessness on the one hand and negligence on the other. But a construal of awareness as *occurrent* might unduly restrict the scope of knowledge and recklessness, and in any event gives rise to difficulties in specifying the time at which the defendant must have had the occurrent state.

silent on the matter, current jury instructions leave the issue up to whatever type of belief, awareness, etc., the case happens to have primed. At the very least, the type that is primed is not always the type that commentators advocate employing or presume is already being employed in criminal cases.<sup>205</sup> Moreover, although more research must be done to determine more precisely what sorts of instruction wording and factual circumstance prime given types of belief, the studies surveyed in Part III evince patterns that, even at this early stage, make some instances of unintended priming predictable and potentially addressable.<sup>206</sup>

#### *D. The Vanishing Distinction Between Description and Judgment*

One broader divergence between mens rea theory and practice warrants mention. An oft-recited piece of conventional wisdom in criminal law theory is that “[m]odern criminal law codes . . . tend to make greater use of purely ‘descriptive’ criteria, relative to ‘evaluative’ criteria,” in their mens rea schemas than did older criminal codes.<sup>207</sup> As Alan C. Michaels explains, “descriptive” mens rea standards “identify the grounds for liability and include those grounds in terms that do not require normative judgment for

---

205. See *supra* Part II. For example, as a general matter it seems unlikely that judges, commentators, or legislators would endorse a thick or occurrent construal of belief, awareness, conscious disregard, etc., in any case in which a jury would thereby find that the defendant was *not* reckless (because he lacked the relevant type of belief) but *was* knowing (because he possessed some other type of belief). See *supra* notes 101-02 and accompanying text (noting the supposed incoherence of a finding of knowledge that P without belief that P); *supra* note 12 and accompanying text (noting the assumption that greater culpability attaches to knowledge than to “mere” recklessness).

206. See *supra* Section III.B.1.b. It is worth stressing that here, as elsewhere in this Article, one possible alternative normative response to jury instruction ambiguity would be to embrace it. Perhaps in each case jurors will tend to settle on the meaning of the term at issue that, given the specific facts of the case, more closely tracks culpability than would more precise mental state categories devised *ex ante*. This sort of response, familiar in debates over the desirable degree of specificity in criminal law’s prohibitions, runs into the problems of fair notice, etc., noted *supra* note 202. It is also problematic insofar as the jury lacks information about, or ability to tailor, the defendant’s likely punishment, and is instead required to make a binary guilt or innocence decision concerning each charged offense.

207. Simons, *Understanding the Topography*, *supra* note 42, at 246; see also R.A. Duff & Stuart P. Green, *Introduction to DEFINING CRIMES* 10-16 (R.A. Duff & Stuart P. Green eds., 2005) (distinguishing the “descriptivist” from the “moralist” approach); PILLSBURY, *supra* note 42, at 83-85 (contrasting the “allusive style of mens rea” typified by English common law but represented as well in the MPC’s conception of recklessness, with the “analytic style” of mens rea found in other MPC belief state concepts); GEORGE P. FLETCHER, *RETHINKING CRIMINAL LAW* 396-400 (1978).

their application.”<sup>208</sup> In contrast, evaluative, or “judgmental,” mens rea standards “define criminal liability in plainly indeterminate terms that call for appraisals, assessments, or judgments beyond findings of fact.”<sup>209</sup> Within the MPC, for example, recklessness and negligence are thought to retain some of the common law’s reliance on evaluation,<sup>210</sup> while “[p]urpose and knowledge are fully descriptive,” as is the “‘substantial risk’ component of the recklessness criterion.”<sup>211</sup>

This conventional account of a modern trend toward description and away from evaluation is somewhat misleading. “Knowledge,” for example, held up as a paradigmatic example of an objective, value-neutral, and a-contextual belief state on both legal and philosophical accounts,<sup>212</sup> turns out instead to be like recklessness: the very same mental events can be knowledge or not knowledge depending on the moral valence of the actions to which they are relevant and the pragmatic context in which the action takes place. The same may be said for other supposedly non-evaluative mental states invoked in current jury instructions.<sup>213</sup>

The point here is not that modern mens rea remains evaluative because jurors engage in blame-based motivated reasoning that gets in the way of their following directions. Quite to the contrary, the point is that modern criminal codes, and jury instructions based on them, do not instruct jurors to employ non-evaluative concepts. Instead, current instructions employ concepts like “knowledge,” and descriptions like “high probability” that, on

---

208. Alan C. Michaels, *"Rationales" of Criminal Law Then and Now: For a Judgmental Descriptivism*, 100 COLUM. L. REV. 54, 62 (2000).

209. *Id.* at 64.

210. See, e.g., PILLSBURY, *supra* note 42, at 83-85; Simons, *Should the Model*, *supra* note 7, at 199 (“The MPC’s current provisions are a mix of descriptive and more evaluative criteria. Purpose and knowledge are fully descriptive” as is “[t]he ‘conscious . . . of a substantial risk’ component of the recklessness criterion”).

211. Simons, *Should the Model*, *supra* note 7, at 199; see also Simons, *Understanding the Topography*, *supra* note 42, at 248 n.51 (arguing that Alexander and Ferzan “greatly overstate[] the extent to which evaluative judgments by juries actually affect criminal liability,” since “[m]any crimes contain mens rea requirements of knowledge or purpose[,]” which are not evaluative concepts) (citing ALEXANDER & FERZAN, *supra* note 14, at 292). Professor Ferzan argues that the availability of affirmative defenses (excuses and justifications) insert evaluation into what would otherwise be, in the case of purpose or knowledge, for example, a purely descriptive inquiry. See Ferzan, *supra* note 42, at 2536, 2529-30; see also Alexander, *supra* note 42, at 940; ALEXANDER & FERZAN, *supra* note 14, at 32-33. In contrast, this Article emphasizes the degree to which evaluation is part of the mens rea concepts at issue in the prima facie mens rea requirement.

212. See, e.g., sources cited *supra* notes 25, 42, 207.

213. See sources cited *supra* notes 149, 211.

their most natural lay-interpretation, turn out to be both descriptive *and* evaluative.<sup>214</sup>

*E. The Need for Greater Scrutiny of Mens Rea Jury Instructions*

Before concluding, this Article highlights three potential reasons why mens rea jury instructions and their interpretation have received relatively little attention to date, along with explanations for why each reason is misguided.

First, even if one is convinced that problematic ambiguities lurk within mens rea jury instructions, one might still think that careful scrutiny of jury instructions is unimportant because jurors are too biased or incompetent to respond to changes in jury instructions. This objection is overblown. The studies reviewed in Part III, along with numerous other social scientific studies, show that jury-eligible laypeople's belief-state ascriptions are responsive to small changes in wording, and that their belief-state ascriptions are largely constrained by instructions and not simply the product of irrational, "blame early"-style motivated reasoning.<sup>215</sup> Moreover, even though irrational bias likely influences mental-state ascription to some degree, alteration of misleading jury instructions represents a simpler and more immediate reform than the sort of "de-biasing" efforts sometimes gestured at in the legal literature on criminal law and psychology.<sup>216</sup>

One might nonetheless suggest a second set of reasons for why scrutiny of mens rea jury instructions is not so pressing: perhaps legal institutions already ensure reasonably well-functioning and ever-improving jury

---

214. This point is important not only for accurately describing current law, but also for determining whether or how to implement various proposals for criminal law reform. *See, e.g.,* Youngjae Lee, *Reasonable Doubt and Moral Elements*, 105 J. CRIM. L. & CRIMINOLOGY (forthcoming 2016) (arguing that "the beyond a reasonable doubt requirement should not apply to moral or normative elements" of crimes but should continue to apply to purely descriptive elements of crimes).

215. Of course, criminal trials differ in important respects from social science experiments like the ones described in this Article, which, for example, do not involve group deliberation. This does not mean, however, that they fail to predict juror behavior. *See generally* Jonathan J. Koehler & John B. Meixner, *Jury Simulation Goals*, in *THE PSYCHOLOGY OF JURIES: CURRENT KNOWLEDGE AND A RESEARCH AGENDA FOR THE FUTURE* (Margaret Bull Kovera ed., forthcoming 2016).

216. *See, e.g.,* Nadelhoffer, *Bad Acts*, *supra* note 2, at 211-12 (suggesting that "[p]erhaps, if jurors were made aware of the various—and seemingly predictable—ways that their judgments can be unwittingly affected by evaluative considerations and blame-validation biasing, they would be better able to live up to their legal duty to base their decisions solely on the material facts of the case," though noting that such reform would face serious difficulties).

instructions without the need for an account of laypeople's belief-state ascription practices. After all, confused jurors can ask clarifying questions. Lawyers can propose alternative instructions if they are concerned that jurors' implicit understanding of belief state terms will hurt their client. Appellate judges can correct and refine over time the instructions lower courts provide. More generally, if jury instructions were really giving a systematically different impression than judges and commentators thought they gave, then the resulting false negatives and false positives in jury findings would function as "red flags," alerting us to the problem.<sup>217</sup> Perhaps the absence of such red flags in the mens rea context means there is no problem in the first place.

Unfortunately, current institutional mechanisms do not warrant such an optimistic view. Jurors might not consciously consider the ways they are resolving ambiguities in mental-state descriptions, and they are even less likely to note how their resolutions of such ambiguities depart from those of legal commentators.<sup>218</sup> Lawyers, insofar as they are aware of the sorts of subtle ambiguities noted in Part III, are unlikely to successfully propose instructions that stray far from typical instructions, and even if they do, the resulting instruction will in no way bind, or likely even influence, other judges' instructions. Moreover, even systematically "inaccurate" mens rea ascription will not raise red flags; it is exceedingly difficult to show that a jury in any given case got the actus reus finding right but the mens rea finding wrong. Hence no evidence of false positives—no red flags—will draw attention to systematically problematic mens rea instructions, even as studies of exonerations reveal ways actus reus determinations can go systematically awry.

As for appellate review, it ensures that jury instructions do not misstate the law,<sup>219</sup> but it does not ensure optimal jury instructions,<sup>220</sup> and in fact

---

217. A parallel might be drawn to the way false positive guilty verdicts have led to critical examination of pre-trial line-ups and interrogations. Cf. Robert P. Burns, *Some Limitations of Experimental Psychologists' Criticisms of the American Trial*, 90 CHI.-KENT L. REV. 899 (2015).

218. Nor could jurors likely craft an effective clarifying question, especially given judges' understandable hesitancy to wade into thorny and unsettled legal issues in response to a jury's mid-deliberation questions.

219. That said, where counsel fails to lodge an objection to a given instruction, a misstatement of the law will only be overturned on appeal if it was not harmless. See, e.g., *United States v. Khan*, 53 F.3d 507, 516-17 (2d Cir. 1995).

220. Indeed, appellate courts sometimes strongly support the use of particular wording in jury instructions while nonetheless not overturning convictions based on instructions that failed to include it, even without engaging in harmless error analysis. For example, in *Jewell*,



further ensures that instructions will remain minimally error-prone and hence less than fully explanatory.<sup>221</sup> For better or worse, the process of jury-instruction creation and review ensures cautious instructions that are highly path-dependent—an amalgam of statements that happen to have been reviewed and not found erroneous by the relevant appellate court.

In short, current institutional mechanisms are ill-designed to note and remedy problematic mens rea jury instructions, suggesting all the more strongly the need for further research in this area. For now, the upshot of the current system of jury-instruction creation and review is to put pressure on legislatures. Their definition of a given mental state must be aimed not only at judges but also at jurors, who will often receive minimal-to-no guidance beyond the statutory term itself and any statutory definitions.

A third and final objection to this Article's focus on jury instructions is more theoretical. Perhaps, in setting out to understand what belief state concepts "mean" as they are used in criminal mens rea analysis, this Article's focus on jury instructions and their interpretation is incomplete or misleading. After all, jury instructions can get the law wrong, and even those that get the law right can still be misinterpreted, maybe even systematically so. The very *possibility* of incorrect instructions or mistaken interpretation implies that there is some "legal meaning" of these mens rea concepts that can differ from what reasonable jurors, faithfully applying instructions, would understand these terms to mean.

This objection highlights an interesting and under-theorized issue: How do jurors and jury instructions fit into debates about the content of the law, as well as the relation between the content of the law and the

---

the Ninth Circuit upheld a willful ignorance instruction, stressing the importance of the phrase "solely and entirely a result of"—as used in the phrase "[you may convict if his] ignorance . . . was solely and entirely the result of . . . a conscious purpose to disregard" the possibility that he was transporting narcotics—as a bulwark against a conviction for mere recklessness. *United States v. Jewell*, 532 F.2d 697, 704 n.21 (9th Cir. 1976) (en banc). But, as Judge Sand notes in his influential treatise on federal jury instructions, after *Jewell* it is still "not clear whether this language is ever required, even in the Ninth Circuit," which appears never to have reversed a conviction "for the failure to include it." SAND ET AL., *supra* note 50, at 7. Similarly, in articulating the proper mens rea for aiding and abetting in *Rosemond*, the Supreme Court at times invoked the defendant's "full knowledge" and "full awareness," but there is no indication that these phrases were meant to, or will, show up in future jury instructions concerning aiding and abetting. *See Rosemond v. United States*, 134 S. Ct. 1240, 1248-50 (2014).

221. *Cf.* *United States v. Aguilar*, 80 F.3d 329, 331 (9th Cir. 1996) (warning that on review, the appellate court will "carefully" review any attempt at defining "knowledge," and that "[a]lthough a correct instruction may assist jurors in understanding knowledge," "[t]he district court is not necessarily required to define knowledge").

communicative content of legal texts?<sup>222</sup> Delving into that issue would, however, take us far afield. For this Article's purposes, it is enough to note the ways in which mens rea belief state concepts, as they are understood and applied by fact-finders in criminal trials, differ from the concepts that judges presume are operative, and that scholars use in their descriptive accounts and normative proposals concerning mens rea. Which account of mens rea belief-state concepts might more properly be labeled an account of the "true legal meaning" of those terms is, for present purposes, beside the point.

### *V. Conclusion*

This Article began with a simple question: What do legally relevant belief states consist of? Focusing on criminal law, its method of answering that question has been to examine jury instructions and jurors' likely interpretation and application of them in the face of circumstantial evidence of mental state.<sup>223</sup> Drawing on empirical research previously overlooked in the legal literature, this Article helps fill a gap in our understanding of legally relevant belief states, revealing several important divergences between theory and practice.

As a general matter, belief-state ascription is much more responsive to the perceived practical and normative aspects of an agent's situation than has traditionally been assumed. The amount of evidence one must have in order to be taken to have formed a belief that P, and the amount of certainty needed to make that belief knowledge, depend on what courses of action were available to the agent, as well as how counter-normative the agent's P-relevant action was. Moreover, concepts like belief, awareness, and conscious disregard can mean quite different things in different contexts. As a result, supposedly legally irrelevant factual differences and slight alterations in jury-instruction wording systematically prime jurors to

---

222. Cf. Lawrence B. Solum, *Communicative Content and Legal Content*, 89 NOTRE DAME L. REV. 479, 479 (2013) (drawing a distinction between "communicative content" and "legal content," arguing that the relationship between the two "varies with context; different kinds of legal texts produce different relationships between linguistic meaning and legal rules," and discussing constitutions, statutes, judicial opinions, and contracts, but not jury instructions).

223. Though outside the scope of this Article, a similar approach has interesting implications for other areas of law in which belief state ascription plays an important role, such as criminal procedure, torts, contracts, and statutory interpretation.

consider occurrent or dispositional and “thick” or “thin” forms of the concepts in question, ultimately altering case outcomes.

As an initial step toward understanding what belief-state terms mean in practice, this Article should help clarify what is at stake in, and how best to implement, various proposals for mens rea reform. Its overarching implication is that jury instructions deserve closer scrutiny and empirical testing than has previously been undertaken. Growing appreciation for the biases that infect the criminal trial process should not blind us to the possibility that relatively minor adjustments to jury instructions could positively affect outcomes in a multitude of criminal cases, bringing mens rea concepts more in line not only with what we have assumed they do mean, but also with what we suggest they should mean.