Hastings Law Journal

Volume 73 | Issue 4 Article 2

5-2022

Untangling Right from Wrong in Insanity Law: Of Dogs, Wolves & God

Kate E. Bloch

Articles

Untangling Right from Wrong in Insanity Law: Of Dogs, Wolves & God

KATE E. BLOCH[†]

In almost all U.S. jurisdictions, a qualifying mental illness that prevents an accused from distinguishing right from wrong can provide support for a determination of legal insanity. Nonetheless, "wrongfulness" remains a term of myriad and somewhat elusive meanings. Instead of enhancing clarity, by engaging with simplified examples, the U.S. Supreme Court's broad-brush approach in its 2020 majority and dissenting opinions in Kahler v. Kansas *threatens to exacerbate confusion about "wrongfulness" in legal insanity doctrine. This Article surfaces challenges in the Court's and dissent's analyses and aims to discourage reliance on problematic assumptions about "wrongfulness" in insanity law.*

[†] Professor of Law and Harry & Lillian Hastings Research Chair, University of California, Hastings College of the Law. I am grateful to colleagues who provided invaluable insights on earlier versions of this Article, especially Evan Lee, Jessica Vapnek, and Lois Weithorn. I am also indebted to forensic psychiatrist Dr. Jeffrey Gould (co-author on a previous article) for his wisdom and guidance on an earlier draft of this Article. I benefitted from his decades of experience evaluating individuals advancing insanity claims in criminal cases. I am also grateful for the fine work of my library liaison, Vince Moyer, and my research assistant, Richard Yates.

TABLE OF CONTENTS

INTRODUCTION

If, "as a result of severe mental illness, the defendant thought that a dog ordered him to kill"[1] a human being, and the defendant killed someone in response to that command, should that qualify as legal insanity? Justice Breyer opened Part I of his March 2020 dissent in the U.S. Supreme Court decision in *Kahler v. Kansas*[2] with this "much-simplified example."[3] He then asserted that "[u]nder the insanity defense as traditionally understood, the government cannot convict [the] defendant"[4] in these circumstances. Justice Breyer introduced this example to highlight the fundamental importance of moral blameworthiness to liability in criminal insanity jurisprudence.[5] In Justice Breyer's description, because the accused, as a result of severe mental illness, thought that a dog ordered him to kill the victim, the defendant did not know what he was doing was wrong; therefore, such blameworthiness was absent, and a criminal conviction should not follow.[6]

Justice Kagan's majority opinion in *Kahler* referenced the same example.[7] The associated discussion implied that, if the defendant believed their acts were morally justified, insanity tests that include a failure to understand the moral wrongfulness of the conduct would allow the defense to prevail, whether the commander were God or a dog, as in the dissent's example.[8]

At least since 1843, assuming that the conduct results from a qualifying mental illness, insanity doctrine has provided potential legal recourse for a

1. Kahler v. Kansas, 140 S. Ct. 1021, 1038 (2020) (Breyer, J., dissenting). Justice Breyer offered a similar, but not identical, example in a dissent to a denial of certiorari in *Delling v. Idaho*, 568 U.S. 1038 (2012). The example from the *Delling* case is discussed *infra* notes 132–39 and accompanying text.

2. 140 S. Ct. at 1038. In the dissent, the example was presented in declarative sentences rather than presented in the form of a question as it appears here.

3. *Id.* Justice Breyer was joined in the dissent by Justices Ginsburg and Sotomayor.

4. *Id.*

5. *Id.* ("Kansas has not simply redefined the insanity defense. Rather, it has eliminated the core of a defense that has existed for centuries: that the defendant, *due to mental illness*, lacked the mental capacity necessary for his conduct to be considered morally blameworthy."). The example was one of a pair of examples upon which Justice Breyer relied. *Id.* The example in the text focused on the moral incapacity prong of traditional insanity tests. *Id.* The other example, where the defendant shoots a person believing the person is actually a dog, related to a cognitive incapacity prong, which Kansas retained in its insanity approach. *Id.*

6. *Id.* at 1038, 1048 ("[T]he defendant . . . differ[s] from ordinary persons in ways that would lead most of us to say that they should not be held morally responsible for their acts."). The *Kahler* majority also suggests that the dissent's example includes a recognition that the defendant believes the response to the dog's command to be "morally justified[.]" *Id.* at 1036 (majority opinion). In addition, Justice Breyer writes that the dog-as-commander example "falls within [*M'Naghten's*] second prong." *Id.* at 1038 (Breyer, J., dissenting). This is the prong relating to whether the defendant knew that their conduct was wrong. Queen v. M'Naghten, (1843) 8 Eng. Rep. 718, 722 (HL). Moreover, because Justice Breyer is using the example to demonstrate the difficulty with the Kansas approach, one might infer that the dissent intends that the defendant in the example does not know their conduct to be wrong.

7. *Id.* at 1036 (majority opinion). Justice Kagan's majority opinion was joined by Chief Justice Roberts and Justices Thomas, Alito, Gorsuch, and Kavanaugh.

8. *See id.* ("And the defendants they will treat differently are exactly those Kahler (and the dissent) focus on: those who know exactly what they are doing (including that it is against the law) but believe it morally justified—because, say, it is commanded by God (or in the dissent's case, a dog).") (citation omitted).

defendant who did not, or was unable to, understand that the act they performed was wrong.[9] The seminal 1843 *M'Naghten* decision defines the wrongfulness prong of legal insanity as: "at the time of the committing of the act, the party accused was labouring under such a defect of reason, from disease of the mind . . . that he did not know he was doing what was wrong."[10] Consistent with *M'Naghten*, approximately 90% of U.S. jurisdictions[11] maintain some version

9. *See* Clark v. Arizona, 548 U.S. 735, 747–49 (2006); *M'Naghten*, 8 Eng. Rep. at 722. Pursuant to *M'Naghten*, defendants can, in the alternative, show that they did not know the nature and quality of their act to invoke legal insanity. *Id.* Justice Kagan's majority opinion notes, however, that "[o]nly with *M'Naghten*, in 1843, did a court articulate, and momentum grow toward accepting, an insanity defense based independently on moral incapacity." *Kahler*, 140 S. Ct. at 1034. *But see id.* at 1040–45 (Breyer, J., dissenting).

10. *M'Naghten*, 8 Eng. Rep. at 722. Some scholars parse the *M'Naghten* case as articulating both a general right and wrong test and a separate more specific rule that applies to delusions. *See, e.g.*, E. Lea Johnston & Vincent T. Leahey, *The Status and Legitimacy of* M'Naghten's *Insane Delusion Rule*, 54 UC DAVIS L. REV. 1777, 1786 (2021) (noting that "whether the insane delusion rule merely provides an example of the general right-and-wrong test, or whether it creates a distinct test for irresponsibility that could supplement the general insanity standard or even detract from it" has provoked differing views) (footnote omitted)). For a detailed analysis of this question and its impact, see generally Johnston & Leahey, *supra* note 10. For purposes of this Article, the analysis treats *M'Naghten's* commentary on partial delusions and the need to treat the facts as the defendant perceived them as applicable here. This Article does not advance a normative view of how wrongfulness should be defined nor of how legal insanity more generally should be codified.

11. The reference to U.S. jurisdictions in this Article encompasses those jurisdictions whose approaches to insanity are described in the Appendix to *Kahler*, 140 S. Ct. at 1051–59, namely the approaches in the 50 states, the federal government, and the District of Columbia. The approximately 90% statistic was calculated from the dissent and Appendix in *Kahler*, in which 47 out of the 52 described jurisdictions appear to offer a blameworthiness evaluation option. 140 S. Ct. at 1046. The majority opinion, however, characterizes the 2012 North Dakota statutory definition as "replacing the right-from-wrong test with an inquiry into whether the defendant's act arose from '[a] serious distortion of [his] capacity to recognize reality[.]'" *Id.* at 1035 (majority opinion). If one adopts the majority's assignment of North Dakota, then that would change the statistic to 88.46% (46 of 52 jurisdictions). The dissent explains:

> Today, 45 States, the Federal Government, and the District of Columbia continue to recognize an insanity defense that retains some inquiry into the blameworthiness of the accused. Seventeen States and the Federal Government use variants of the M'Naghten test, with its alternative cognitive and moral incapacity prongs. Three States have adopted M'Naghten plus the volitional test. Ten States recognize a defense based on moral incapacity alone. Thirteen States and the District of Columbia have adopted variants of the Model Penal Code test, which combines volitional incapacity with an expanded version of moral incapacity. See Appendix, *infra*. New Hampshire alone continues to use the "product" test, asking whether "a mental disease or defect caused the charged conduct." State v. Fichera, 153 N. H. 588, 593, 903 A. 2d 1030, 1035 (2006). This broad test encompasses "'whether the defendant knew the difference between right and wrong and whether the defendant acted impulsively,'" as well as "'whether the defendant was suffering from delusions or hallucinations.'" State v. Cegelis, 138 N. H. 249, 255, 638 A. 2d 783, 786 (1994). And North Dakota uses a unique formulation that asks whether the defendant "lacks substantial capacity to comprehend the harmful nature or consequences of the conduct, or the conduct is the result of a loss or serious distortion of the individual's capacity to recognize reality." N. D. Cent. Code Ann. § 12.1–04.1–01(1) (2012).
>
> Of the States that have adopted the M'Naghten or Model Penal Code tests, some interpret knowledge of wrongfulness to refer to moral wrong, whereas others hold that it means legal wrong."). See also Delling v. Idaho, ("The law has long recognized that criminal punishment is not appropriate for those who, by reason of insanity, cannot tell right from wrong. See 4 W. Blackstone, Commentaries on the Laws of England 24–25 (1769); [Queen v. M'Naghten, (1843) 8 Eng. Rep. 718 (HL)]. The insanity defense in nearly every State incorporates this principle. See Clark v. Arizona, 548 U.S. 735, 750–752 (2006) (noting that all but four States recognize some version of the insanity

of this blameworthiness inquiry[12] as a recognized excuse or defense to a criminal conviction.[13] Pursuant to this test, if, due to a qualifying mental illness, the accused engaged in conduct without an ability to understand or understanding that the act was wrong,[14] as defined in that jurisdiction, the test could excuse the defendant.[15] In contrast, engaging in such conduct when the accused had an ability to understand or understood that the act was indeed wrong,[16] generally undermines an insanity excuse under this test.[17]

Justice Breyer describes the blameworthiness test as asking: "even if the defendant knew what he was doing, did he have the capacity to know that it was wrong?"[18] But, then, without analysis of that wrongfulness issue in the dog-as-

---

defense); [R. BONNIE, A. COUGHLIN, J. JEFFRIES, & P. LOW, CRIMINAL LAW 604 (3d ed. 2010)] (same).

Id. at 1046 (Breyer, J., dissenting); *see also* Delling v. Idaho, 568 U.S. 1038, 1038 (2012) (Breyer, J., dissenting in denial of certiorari) ("The law has long recognized that criminal punishment is not appropriate for those who, by reason of insanity, cannot tell right from wrong.") (citations omitted).

12. Formulations differ on the precise terminology regarding awareness of wrongfulness. *See, e.g.*, E. Lea Johnston, *Delusions, Moral Incapacity, and the Case for Moral Wrongfulness*, 97 IND. L.J. 297, 308–16 (2022) (exploring formulations at length, as well as much recent cognitive science research, and, overall, arguing for a broader conception of wrongfulness in insanity law). Professor Johnston opines that "forensic mental health examiners should consider the systemic, cognitive, and affective distortions associated with delusions, in addition to delusions' faulty content, in sanity evaluations and assess the extent they could have undermined a defendant's ability to understand the wrongfulness of her act." *Id.* at 329. As noted above, *M'Naghten* uses "know"; the Model Penal Code uses "substantial capacity to appreciate." *M'Naghten*, 8 Eng. Rep. at 722; MODEL PENAL CODE § 4.01 (AM. L. INST. 1985). California refers to the defendant being "incapable of knowing or understanding . . . or distinguishing right from wrong." CAL. PENAL CODE § 25(b) (2020). Although the specific formulation is important in application to real cases, that question is not central to the analysis here. To avoid confusion, this Article refers primarily to "an ability or inability to understand."

13. *Kahler*, 140 S. Ct. at 1046 (Breyer, J., dissenting).

14. Or a lack of substantial capacity to comprehend the harmful nature or consequences under the North Dakota terminology. *See* N.D. CENT. CODE ANN. § 12.1-04.1-01(1) (West 2012).

15. *See, e.g.*, *Clark*, 548 U.S. at 748. However, it is not clear whether the specific jurisdictional definition of legal insanity is necessarily determinative in jury decisionmaking. *See e.g.*, Randy Borum & Solomon M. Fulero, *Empirical Research on the Insanity Defense and Attempted Reforms: Evidence Toward Informed Policy*, 23 L. & HUM. BEHAV. 375, 380 (1999) (describing research related, inter alia, to potential limitations in juror comprehension of insanity instructions as well as conflicting evidence about the impact of actual changes in insanity standards on jury decisionmaking).

16. As defined in that jurisdiction.

17. In the New Hampshire evaluation, whether the accused understood or could understand that the act was wrong is only one of the factors a juror can evaluate in determining sanity. *See* State v. Cegelis, 638 A. 2d 783, 786 (1994). Consequently, perceiving the act to be morally unjustified would not necessarily lead to a sanity finding. Moreover, not all jurisdictions focus on the defendant's personal understanding of whether the conduct was wrongful. Instead, a jurisdiction may substitute a societal evaluation of whether the conduct was wrongful from the perspective of a reasonable person using the facts of the events as the defendant understood them. For a discussion of the approach of such a jurisdiction, see *infra* note 77 and accompanying text.

18. *Kahler*, 140 S. Ct. at 1038–50. Justice Breyer does indicate that the *M'Naghten* formulation is not the only possible one to describe the blameworthiness inquiry. *Id.* at 1039. He notes specifically that he "do[es] not mean to suggest that *M'Naghten*'s particular approach to insanity is constitutionally required . . . *M'Naghten*'s second prong is merely one way of describing something more fundamental. Its basic insight is that mental illness may so impair a person's mental capacities as to render him no more responsible for his actions than a young child or a wild animal." *Id.* Justice Breyer explains further that "[a]lthough English and early American sources differ in their linguistic formulations of the legal test for insanity, with striking consistency, they all

commander example, in the very next sentence, Justice Breyer summarily announces that "[a]pplying this test to my example [of the dog commanding the killing], a court would find that [the defendant] successfully established an insanity defense."[19] As a result, the dissent effectively equates the defendant's illness-generated command hallucination with an inability to distinguish right from wrong. While some forms of mental illness "can seriously impair a sufferer's ability rationally to appreciate the wrongfulness of conduct[,]"[20] research on mental illness also suggests that command hallucinations can sometimes be effectively ignored or resisted by the person experiencing them.[21] Consequently, the fact that someone suffers from command hallucinations does not necessarily mean that they could not act within the confines of the law or distinguish right from wrong at the time of the specific conduct.[22]

Professor Steven R. Smith also flags the problematic nature of Justice Breyer's example in a brief mention in an article reviewing the *Kahler* decision when Professor Smith characterizes the dog example as "[a]n unfortunate" one that "may somewhat confuse the issues…"[23] He opines that "[a] defendant

express the same underlying idea: A defendant who, due to mental illness, lacks sufficient mental capacity to be held morally responsible for his actions cannot be found guilty of a crime." *Id.*

19. *Id.* at 1038.

20. Brief of American Psychiatric Ass'n. & American Academy of Psychiatry and the Law as Amici Curiae in Support of Petitioner at 15, Delling v. Idaho, 568 U.S. 1038 (2012) (No. 11-1515) ("Psychiatrists' clinical experience, as well as the peer-reviewed research literature, support the conclusion that severe mental illness can seriously impair a sufferer's ability rationally to appreciate the wrongfulness of conduct."); *see also Delling*, 568 U.S. 1038 (Breyer, J., dissenting in denial of certiorari). For additional interpretations of Justice Breyer's example, see, for instance, *infra* notes 23–24 and accompanying text.

21. *See, e.g.*, Paul Chadwick & Max Birchwood, *The Omnipotence of Voices: A Cognitive Approach to Auditory Hallucinations*, 164 BRIT. J. PSYCHIATRY 190, 190–95, 197, 200 (1994) ("In our group, the severity of the command, and not beliefs, was the single most important determinant of compliance—there was no compliance with life-threatening commands, and compliance with mild commands was commonplace."). Chadwick and Birchwood do note that "the belief that a voice comes from a powerful and vengeful spirit may make the person terrified of the voice and comply with its commands to harm others; however, if the same voice were construed as self-generated, the behaviour and affect might be quite different." *Id.* at 190–91. Their study focused on engaging participants in perceiving the voice as self-generated. *Id.* at 196.

22. *See id.*

23. Steven R. Smith, *Supreme Court 2019–2020: Insanity, Discrimination, and DACA—and a Pandemic*, 46 J. HEALTH SERV. PSYCH. 181, 182 (2020). In accompanying footnote 10, Professor Smith describes the concern as follows:

> The dissent raises an example of why the moral capacity branch matters. It imagines two defendants, both charged with murder. They both have severe mental illness. The mental illness causes the first defendant to think "the victim was a dog;" it causes the second to think "that a dog ordered him to kill the victim. Under the insanity defense as traditionally understood, the government cannot convict either defendant. Under Kansas' rule, it can convict the second but not the first." There are not additional facts given in the hypothetical.
>
> It is clear in the first example that wrongly believing the person is a dog would be a defense because the defendant would not understand, given the delusion, that he is killing a person. As to the second defendant, however, it is not so clear (absent unstated facts). The question essentially is, why would a delusion that a dog told a defendant to kill a person implicate the moral capacity defense. How would he be morally justified in killing the victim even if the dog had, in fact, ordered him to do so?

killing what he believes to be a person cannot legitimately (without more facts) think he is morally or legally justified in doing so just because a dog ordered him to do so."[24]

In the majority's version of the God and dog examples, the opinion explicitly asserts that the defendant believed that the conduct was "morally justified[.]"[25] This provides express recognition that the defendant entertained a perception with respect to the rectitude of their conduct that might play a role in excusing it.[26] Although not particular to a command hallucination, the opinion even offers a scenario in which such a belief could excuse the defendant.[27] However, by incorporating the dissent's example of the dog's command by reference and treating it in that reference as if it were equivalent for purposes of a wrongfulness evaluation to such a command by God, this Article contends that the majority paints with too broad a brush.

To be fair, Justice Breyer's dog-as-commander example is intended, as he describes it, to be a "much simplified" one.[28] It is one of two examples involving a dog designed to provide a "conceptual"[29] distinction between two prongs of a traditional insanity test and to support Justice Breyer's contention that the Kansas test at issue in *Kahler* is missing the moral blameworthiness prong of that test.[30] Relatedly, Justice Kagan's majority opinion's brief incorporation by reference of the dog example involves just that, not an extended approval of the dog example, and appears intended to portray another important difference between two types of wrongfulness (moral versus legal).[31]

Id. at 182 n.10 (citations omitted). For related concerns about Justice Breyer's dog-as-commander example, see Kent Scheidegger, *An Insanity Debate Goes to the Dogs*, CRIME & CONSEQUENCES (Mar. 23, 2020), https://www.crimeandconsequences.blog/?p=791 (arguing that the hypothetical "makes no sense . . . . Nothing in Justice Breyer's hypothetical negates the conclusion that Defendant Two knew that it was wrong to kill the person."); Rachel Tollefsrud, *Saving the Insanity Defense: Insight into Personality Disorders and the Necessary Elements of the Test*, 48 MITCHELL HAMLINE L. REV. 372, 399 (2022) ("[T]aking Justice Breyer's example from his dissent in *Kahler*, a person who kills because a dog 'told them to' has still murdered even though their reality is significantly skewed. In other words, this person knows what they are doing, knows it is wrong, and probably is able to stop if they wanted to because the reason they are committing the action—the *why*—is not wrapped up in their volition.") (citing Scheidegger, *supra*) (footnote omitted).

24. *Smith*, *supra* note 23, at 182 n.10.

25. *Kahler*, 140 S. Ct. at 1036 (majority opinion). For the majority's treatment specifically of commands perceived as emanating from God, see *id.* at 1026.

26. *See id*.

27. *Id.* (citing 2 JAMES F. STEPHEN, *Relation of Madness to Crime, in* HISTORY OF THE CRIMINAL LAW OF ENGLAND 124, 149 (1883)).

28. *Kahler*, 140 S. Ct. at 1038 (Breyer, J., dissenting).

29. *Id.*

30. *Id.*

31. *Kahler*, 140 S. Ct. at 1036 (majority opinion). The majority opinion also limits the moral capacity inquiry to definitions that relate specifically to moral as opposed to legal wrong. *See id.* While one could contest the exclusion of legal wrongfulness—which is arguably a variation of the blameworthiness inquiry—from the canon of morality or blameworthiness inquiries, to respect the majority's exclusion, critique of the majority opinion here focuses on models that relate to moral wrongfulness, rather than those that focus on legal wrongfulness. Because the dissent, however, contends that "[w]hile there is, of course, a logical distinction between those interpretations, there is no indication that it makes a meaningful difference in practice. The two

Nonetheless, as this Article argues and aims to illustrate, by using the impoverished example of the dog-as-commander without the necessary critical analysis of wrongfulness, both the majority and the dissent risk exacerbating the confusion surrounding the complex doctrine of wrongfulness in insanity law.

Insanity law, like much of criminal law, lies largely within the purview of local or state legislation and interpretation,[32] but can, nonetheless, be subject to constitutional limits.[33] James Kahler called upon this constitutional oversight power with a due process challenge to Kansas' statutory framework on insanity.[34] "Kahler . . . asked th[e] Court to decide whether the Due Process Clause require[d] States to provide an insanity defense that acquits a defendant who could not 'distinguish right from wrong' when committing his crime—or, otherwise put, whether that Clause require[s] States to adopt the moral-incapacity test from *M'Naghten*."[35] The Court ultimately found that Kansas could remain an outlier; despite its failure to offer a traditional version of the moral incapacity defense, its approach did not violate the Due Process Clause.[36]

To reach a decision on that question, the *Kahler* case produced opinions that examine at length the concept of wrongfulness in insanity and its treatment

---

inquiries are closely related and excuse roughly the same universe of defendants[,]" the critique of the dissent also involves a critique under a legal wrongfulness approach. *Kahler*, 140 S. Ct. at 1046 (Breyer, J., dissenting); *see also infra* note 54 and accompanying text.

32. *Kahler*, 140 S. Ct. at 1028 (majority opinion) ("Within broad limits, . . . 'doctrine[s] of criminal responsibility' must remain 'the province of the States.'") (citation omitted). Federal criminal prosecutions are, of course, subject to congressional definitions of insanity. *See* United States v. Ewing, 494 F.3d 607, 615–16 (7th Cir. 2007) ("Under the Insanity Defense Reform Act of 1984 ("'IDRA'"), it is an affirmative defense to a prosecution for a federal crime if 'at the time of the commission of the acts constituting the offense, the defendant, as a result of a severe mental disease or defect, was unable to appreciate the nature and quality or the wrongfulness of his acts.'") (citation omitted). But federal cases represent a very small percentage of criminal cases nationwide. *Compare* Admin. Off. of the U.S. Cts., *Federal Judicial Caseload Statistics 2020, U.S. District Courts, Criminal Filings*, U.S. CTS, https://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2020 (noting 93,213 federal criminal defendant filings in U.S. district courts in 2020) (last visited Apr. 15, 2022) *with CSP STAT Criminal, Trial Court Caseload Overview, Caseload Detail – Total Criminal, 2020 Incoming Cases*, CT. STATS. PROJECT, https://www.courtstatistics.org/csp-stat-nav-cards-first-row/csp-stat-criminal (last visited Apr. 15, 2022) (noting 9,244,371 state court incoming cases in 32 states in 2020). In addition, commentators note the infrequency with which the insanity defense is successful. *See, e.g.*, Eugene M. Fahey, Laura Groschadl & Brianna Weaver, *"The Angels that Surrounded My Cradle": The History, Evolution, and Application of the Insanity Defense*, 68 BUFF. L. REV. 805, 807–08 (2020) ("In the State of New York, of approximately five thousand murder cases between 2007 and 2016, just six ended with the defendant found not responsible by reason of mental disease or defect; the state does not track how often the defense is raised. Between 2013 and 2017, only eleven defendants out of 19,041 felony and misdemeanor trials conducted in the state were found not responsible by reason of mental disease or defect pursuant to Penal Law § 40.15, and 241 defendants entered an insanity plea out of 1,375,096 convictions during the same time period. As of June 30, 2018, 260 insanity acquittees were receiving treatment in secure confinement and another 452 insanity acquittees 'were in the community subject to orders of conditions.'") (footnotes omitted).

33. *Kahler*, 140 S. Ct. at 1027 ("Under well-settled precedent, a state rule about criminal liability—laying out either the elements of or the defenses to a crime—violates due process only if it 'offends some principle of justice so rooted in the traditions and conscience of our people as to be ranked as fundamental.'") (citation omitted).

34. *Id.*

35. *Id.*

36. *Id.* at 1037.

in various jurisdictions.[37] While these extended analyses of wrongfulness are not the holding of the *Kahler* case,[38] they are, nonetheless, critical for the approximately nine-tenths of jurisdictions nationwide that allow some form of inability to understand wrongfulness to excuse.[39] In these analyses, both the majority and dissent invoke the example of the dog's command to kill.[40]

This Article explores whether the dog's command to commit homicide should qualify as legal insanity under a wrongfulness analysis.[41] The argument

37. *Id.* at 1024–50.

38. The holding in *Kahler* was on the constitutional due process issue, not on the definitions of wrongfulness. *Kahler*, 140 S. Ct. at 1037. However, as, for example, California appellate courts have noted with respect to that state's supreme court, "[e]ven if properly characterized as dictum, statements of the Supreme Court should be considered persuasive." Hubbard v. Superior Court, 66 Cal. App. 4th 1163, 1169 (1997) (quoting United Steelworkers of Am. v. Bd. of Educ. 162 Cal. App. 3d 823, 835 (1984)).

39. *Kahler*, 140 S. Ct. at 1051–59 (Appendix).

40. *See, e.g.*, *Kahler*, 140 S. Ct. at 1036 (majority opinion); *id.* at 1038, 1048 (Breyer, J., dissenting).

41. Justice Breyer analyzes two versions of the dog analogy, one under each of the two prongs of the *M'Naghten* test. *Kahler*, 140 S. Ct. 1021, 1038. The first *M'Naghten* prong involves determining whether the defendant understood the "'nature and quality of the act he was doing . . .'" *Id.* (quoting Lord Chief Judge Tindal in *M'Naghten's Case*, (1843) 8 Eng. Rep. 718, 722 (HL)). The version of the dog analogy that Justice Breyer relates to this first *M'Naghten* prong and the analysis of the first prong itself are not the subject of this Article. The second version, the one provided at the start of this Article, relates to the second *M'Naghten* prong, the wrongfulness question, and is the focus of this Article. For background and easy reference, what follows are excerpts of key pertinent discussions of both versions of the dog analogy and the prongs to which each refers from Justice Breyer's dissent in *Kahler*:

> A much-simplified example will help the reader understand the conceptual distinction that is central to this case. Consider two similar prosecutions for murder. In Prosecution One, the accused person has shot and killed another person. The evidence at trial proves that, as a result of severe mental illness, he thought the victim was a dog. Prosecution Two is similar but for one thing: The evidence at trial proves that, as a result of severe mental illness, the defendant thought that a dog ordered him to kill the victim. Under the insanity defense as traditionally understood, the government cannot convict either defendant. Under Kansas' rule, it can convict the second but not the first.
>
> To put the matter in more explicitly legal terms, consider the most famous statement of the traditional insanity defense, that contained in *M'Naghten's Case*. Lord Chief Justice Tindal, speaking for a majority of the judges of the common-law courts, described the insanity defense as follows:
>
> "[T]o establish a defence on the ground of insanity, it must be clearly proved that, at the time of the committing of the act, the party accused was labouring under such a defect of reason, from disease of the mind, [1] as not to know the nature and quality of the act he was doing; or, [2] if he did know it, that he did not know he was doing what was wrong."
>
> The first prong (sometimes referred to as "cognitive incapacity") asks whether the defendant knew what he was doing. This prong corresponds roughly to the modern concept of *mens rea* for many offenses. The second (sometimes referred to as "moral incapacity") goes further. It asks, even if the defendant knew what he was doing, did he have the capacity to know that it was wrong? Applying this test to my example, a court would find that both defendants successfully established an insanity defense. Prosecution One (he thought the victim was a dog) falls within *M'Naghten*'s first prong, while Prosecution Two (he thought the dog ordered him to do it) falls within its second prong.
>
> In Kansas' early years of statehood, its courts recognized the *M'Naghten* test as the "cardinal rule of responsibility in the criminal law." . . . . But in 1995, Kansas "'legislatively abolish[ed] the insanity defense.'" Under the new provision, a criminal defendant's mental disease or defect is relevant to his guilt or innocence only insofar as it shows that he lacked the intent defined as an element of the offense, or *mens rea.* If the defendant acted with the required level of intent, then he has no defense based on mental illness.

unfolds in two parts. First, the Article furnishes a doctrinal and conceptual overview of wrongfulness in the insanity context.[42] Based on that foundation, the Article evaluates the problematic assertion that obeying a dog's command to kill, one that derives from a qualifying mental illness, even if it involves a defendant's claim of rectitude, per se merits recognition as legal insanity. It concludes that determining whether wrongfulness should excuse depends upon more than whether the defendant suffers from a severe mental illness that generated a command hallucination that caused the conduct and whether they believed they were morally justified. In particular, that determination should involve inquiries about which model or variation of wrongfulness the jurisdiction applies to evaluate insanity claims and why the defendant acted.

I. OVERVIEW OF WRONGFULNESS IN LEGAL INSANITY

With respect to distinguishing right from wrong, Justice Kagan's majority opinion notes that the Supreme Court has "labeled [this prong of] *M'Naghten* . . . a test of 'moral capacity,' and invoked the oft-used phrase 'telling right from wrong' (or in older language, good from evil) to describe its central inquiry."[43] Yet, in legal sanity evaluations, "wrongfulness" remains a term of myriad and sometimes elusive meanings.[44] It lies at the evolving intersection of legal standards of wrongfulness and the complex human dynamics of serious mental illness.[45] Efforts to delineate the genealogy and

---

> Under Kansas' changed law, the defendant in Prosecution One could defend against the charge by arguing that his mental illness prevented him from forming the mental state required for murder (intentional killing of a human being)—just as any defendant may attempt to rebut the State's prima facie case for guilt. The defendant in Prosecution Two has no defense. Because he acted with the requisite level of intent, he must be convicted regardless of any role his mental illness played in his conduct. *See* [State v. Wetrich, 307 Kan. 374, 401, 410 P.3d 105, 125 (2018)] (acknowledging that Kansas' *mens rea* approach "allows conviction of an individual who had no capacity to know that what he or she was doing was wrong").

Id. at 1038–39 (Breyer, J., dissenting) (citations omitted).

> "Now ask, what moral difference exists between the defendants in the two examples? Assuming equivalently convincing evidence of mental illness, I can find none at all. In both cases, the defendants differ from ordinary persons in ways that would lead most of us to say that they should not be held morally responsible for their acts. I cannot find one defendant more responsible than the other. And for centuries, neither has the law."

Id. at 1048.

While *M'Naghten* offered a potential insanity option for both examples, one might argue that shooting a person believing that person to be a dog involves important but different moral (homicide versus animal cruelty) evaluations than if the defendant knew it were a person. In the second example, the moral recognition and evaluation of committing homicide remain, even if one is moved to commit that act through the delusion that the dog commanded it. *See* Smith, *supra* note 23.

42. The models discussed *infra* at notes 60–96 and accompanying text are drawn primarily from Kate E. Bloch & Jeffrey Gould, *Legal Indeterminacy in Insanity Cases: Clarifying Wrongfulness and Applying a Triadic Approach to Forensic Evaluations*, 67 HASTINGS L.J. 913 (2016).

43. *Kahler*, 140 S. Ct. at 1035 (majority opinion).

44. *See e.g.*, State v. Crenshaw, 659 P.2d 488, 491 (Wash. 1983) ("The definition of the term 'wrong' in the *M'Naghten* test has been considered and disputed by many legal scholars.").

45. *See, e.g.*, *Kahler*, 140 S. Ct. at 1028–29, 1037.

applicable models for interpreting wrongfulness in this context populate court opinions and scholarly tomes.[46] As the Supreme Court noted in *Kahler*, even though scholars and jurists refer to this evaluation as one of moral capacity, jurisdictions incorporate more than one doctrinal option for evaluating wrongfulness.[47] Specifically, the Court identified two rubrics for interpretation: moral wrong and legal wrong.[48] Both derive directly or indirectly from at least as far back as the test codified in *M'Naghten*.[49]

Commentators, however, challenge these rubrics on at least two related grounds.[50] First, they ask whether, under the moral wrong approach, consensus morals or norms actually exist in our culturally diverse population today.[51] Second, and sometimes in response to the first criticism, some jurists argue that the two rubrics are, in effect, practically congruent and that distinguishing between moral and legal wrong is unnecessary.[52] To address the criticism of a lack of moral consensus, commentators, including a number of courts, opine that "most cases in which the insanity defense is raised involve crimes sufficiently serious such that society's moral judgment regarding the accused's conduct will be identical to the legal standard reflected in the applicable criminal statute."[53] Or, as Justice Breyer argued in the *Kahler* dissent: "While there is, of course, a logical distinction between those interpretations [moral versus legal wrong], there is no indication that it makes a meaningful difference in practice. The two inquiries are closely related and excuse roughly the same universe of defendants."[54] In effect, if they were almost identical, then legal codes could

46. *See, e.g.*, Brandon A. Yakush & Melinda Wolbransky, *Insanity and the Definition of Wrongfulness in California*, 13 J. FORENSIC PSYCH. PRAC. 355 (2013); United States v. Ewing, 494 F.3d 607, 616 (7th Cir. 2007) ("In the context of the insanity defense, courts and scholars have generally proposed three alternative definitions for the term: (1) legal wrongfulness, as in 'contrary to law'; (2) moral wrongfulness, as in 'contrary to public morality,' determined objectively by reference to society's condemnation of the act as morally wrong; or (3) moral wrongfulness, as in 'contrary to *personal* morality,' determined subjectively by reference to the defendant's belief that his action was morally justified (even if he appreciated that it was illegal or contrary to public morality)."). For a more general condemnation of the state of insanity law and a proposal for a somewhat different role for mental illness in criminal responsibility, see Christopher Slobogin, *An End to Insanity: Recasting the Role of Mental Disability in Criminal Cases*, 86 VA. L. REV. 1199, 1200 (2000) ("[M]ental illness should be relevant in assessing culpability only as warranted by general criminal law doctrines concerning mens rea, self-defense and duress.").

47. *Kahler*, 140 S. Ct. at 1025.

48. *Id.* ("But over the years, 16 States have reoriented the test to focus on the defendant's understanding that his act was *illegal*— that is, legally rather than morally "wrong.").

49. *Id.* The origin of the moral capacity test itself was the subject of much discussion in *Kahler*. *See, e.g.*, *id.* at 1029, 1034; *id.* at 1040–45 (Breyer, J., dissenting).

50. *See, e.g.*, Bloch & Gould, *supra* note 42, at 931.

51. *Id.*

52. *See, e.g.*, *Kahler*, 140 S. Ct. at 1046 (Breyer, J., dissenting).

53. State v. Cole, 755 A.2d 202, 210 (Conn. 2000).

54. *Kahler*, 140 S. Ct. at 1046 (Breyer, J., dissenting). Similarly, later in the dissent, Justice Breyer contends that "the rule adopted by some States that a defendant must be acquitted if he was unable to appreciate the *legal* wrongfulness of his acts, would likely lead to acquittal in the mine run of such cases." *Id.* at 1049 (citation omitted).

supply the societal moral norms, and that would diminish concern about a lack of moral consensus.

Justice Breyer is correct that many cases will produce congruent results, but an important subset of cases will not. Justice Kagan's majority opinion offers an excerpted scenario, drawn from English authority James Fitzjames Stephen, to illustrate where there would likely be a disparity:

> "A kills B knowing that he is killing B, and knowing that it is illegal to kill B, but under an insane delusion that the salvation of the human race will be obtained by . . . the murder of B[.] A's act is a crime if the word 'wrong' [in *M'Naghten*] means illegal. It is not a crime if the word wrong means morally wrong."[55]

Stephen's example, although offering a relatively brief scenario,[56] illustrates the importance of recognizing that the choice between the moral and legal definitions of "wrong" could produce opposite outcomes. In Stephen's 1883 analysis, consistent with doctrine at that time, the legal defense of necessity was not generally recognized as a justification for murder.[57] Under Stephen's analysis (assuming no necessity defense) the murder remains legally wrong, even though, in a moral sphere, the excerpt suggests that the need to save the human race can serve as a moral excuse for the conduct. Thus, Stephen's hypothetical exemplifies the need to recognize both moral and legal wrong rubrics. In addition, the example undergirds the importance of understanding *why* the defendant engaged in the conduct at issue (effort to save the human race). This offers some insight into the defendant's motivation for their conduct and allows us to consider how their conduct relates to morality. This insight and consideration are largely absent from the *Kahler* dissent's dog-as-commander example that inspires this Article.[58]

The moral and legal wrong rubrics lay the foundation for our doctrinal analysis. These are essential starting points. But finer gradations of these rubrics

---

55. *Kahler*, 140 S. Ct. at 1036 (majority opinion) (quoting STEPHEN, *supra* note 27, at 149).

56. In the context of the analysis below, the scenario in the majority opinion might benefit from greater specificity. For instance, the example does not explicitly treat the question of whether the perception of morality is from a societal or personal perspective or both. Nor does the majority's description indicate that the delusion caused A to understand or believe that society would not condemn the act. Interestingly, the majority's rendition represents only an excerpt of Stephen's example. In Stephen's original text, the example read as follows:

> "(3.) A kills B knowing that he is killing B, and knowing that it is illegal "to kill B, but under an insane delusion that the salvation of the human race "will be obtained by his execution for the murder of B, and that God has "commanded him (A) to produce that result by those means. A's act is a "crime if the word 'wrong' means illegal. It is not a crime if the word "wrong means morally wrong.

STEPHEN, *supra* note 27, at 149. The example in its entirety raises the question of implementing a command that the defendant perceives as initiated by God. For a discussion of that consideration more specifically, see *infra* notes 125–127 and accompanying text.

57. *See, e.g.*, Queen v. Dudley & Stephens (1884) 14 QBD 273 (DC).

58. See Justice Breyer's treatment of the example in *Kahler*, 140 S. Ct. at 1038, 1048, 1049 (Breyer, J., dissenting).

inform both the conceptual approaches and real-world applications.[59] Consequently, for purposes of this analysis, the work further divides each of the moral and legal wrong rubrics into two models.[60] The difference between the models within each rubric depends on the extent to which they focus internally on the defendant's world or externally on society's reality, legal codes, or norms. Both sets of models begin with an internal focus where the defendant's perception of the facts of the event, no matter how distorted by their mental illness, governs. As explained in the *M'Naghten* decision, particularly in the context of partial delusions,[61] the defendant "must be considered in the same situation as to responsibility as if the facts with respect to which the delusion exists were real."[62]

In real-world sanity evaluations, the evaluator must assess whether the defendant actually believed the version of events proffered by the defense.[63] The analysis here assumes that the description offered of the defendant's reality is what the defendant actually perceived at the time the relevant acts occurred.[64] It also assumes that the defendant's qualifying mental illness caused the defendant to perceive reality in that way and prompted the defendant's act.[65] These assumptions meet two threshold criteria at the heart of insanity doctrine: that the defendant experienced a qualifying mental illness and that the illness caused the conduct at issue.[66]

Beyond the internally-focused version of the facts common to all these wrongfulness models, the models diverge with respect to the frame for morality and legality. One pair of models focuses on evaluation in a moral context and the other pair on evaluation with respect to the law. Within the moral wrong rubric, the paradigm posits an internal facts-external morality ("external morality") model and an internal facts-internal morality ("internal morality") model. Does the evaluation relate to society's morals[67] or instead to the defendant's own morals? Within the legal wrong rubric, the paradigm posits a relatively analogous pair of models.[68] Does the evaluation depend on society's law or on the defendant's perception of the law?

---

59. *See infra* notes 67–96 and accompanying text.

60. For a more detailed presentation and analysis of these models, see Bloch & Gould, *supra* note 42, at 925–40.

61. *See* Johnston & Leahey, *supra* note 10.

62. Queen v. M'Naghten (1843) 8 Eng. Rep. 718, 722 (HL).

63. Bloch & Gould, *supra* note 42, at 924.

64. For real-world insanity claims, evaluators will generally compare the defendant's account with collateral documentation or other information, which may support or refute the version of events advanced by the defendant in response to the criminal charges.

65. Bloch & Gould, *supra* note 42, at 923.

66. *Id.*

67. As discussed *infra* note 77, and accompanying text, in some external morality applications, the focus is on taking the facts as the defendant understood them and asking if a reasonable person in that situation would have understood the conduct as morally wrongful.

68. This set of models represents a subset of the models proposed in Bloch & Gould, *supra* note 42, at 925–40.

A. MORAL WRONG MODELS

In the moral wrong category, an external morality model commonly focuses on whether the defendant understood that the conduct, based on the facts as perceived by the defendant, would violate society's morals or norms.[69] This anchors the analysis in an external referent, recognizing that there are morals outside the defendant's internally-constructed view against which to measure the "good or evil" of conduct.

For example, a defendant might understand that, in general, the conduct is morally wrong—for instance, the intentional killing of another human being is wrong—but might believe that, if society had the defendant's "special knowledge," society would approve of the conduct. Stephen's excerpted hypothetical could fall into this category. There, the defendant's mental illness caused him to believe that the homicide was necessary for "the salvation of the human race."[70] If, as a result of his special knowledge, the defendant believed that society would approve of his conduct or society deemed his conduct morally acceptable, then this example could lead to a successful invocation of an external morality wrongfulness claim.[71]

Juxtaposed against a successful invocation of an external morality model to find insanity, a defendant might instead personally believe their conduct was morally justified but accurately perceive that society would not approve of their actions. This might arise, for example, when there is no claim of special knowledge.[72] An external morality model does not generally contemplate excusing this conduct.

---

69. Bloch & Gould, *supra* note 42, at 931–36.

70. Kahler v. Kansas, 140 S. Ct. 1021, 1036 (2020) (citing STEPHEN, *supra* note 27, at 149).

71. *Cf.* State v. Wilson, 700 A.2d 633, 640, 643 (Conn. 1997) ("Although we agree with the state that the proper test must incorporate principles of societal morality, we conclude that the state's interpretation of the cognitive prong of § 53a–13 (a) does not sufficiently account for a delusional defendant's own distorted perception of society's moral standards. Accordingly, we conclude that a defendant may establish that he lacked substantial capacity to appreciate the 'wrongfulness' of his conduct if he can prove that, at the time of his criminal act, as a result of mental disease or defect, he substantially misperceived reality and harbored a delusional belief that society, *under the circumstances as the defendant honestly but mistakenly understood them*, would not have morally condemned his actions. . . . The trial court should inform the jury that a person may establish that he was legally insane if he proves that, at the time he committed the prohibited conduct, due to mental disease or defect he suffered from a misperception of reality and, in acting on the basis of that misperception, he did not have the substantial capacity to appreciate that his actions were contrary to societal morality, even though he may have been aware that the conduct in question was criminal. The trial court should instruct the jury further that, in deciding whether the defendant had substantial capacity to appreciate that his conduct was contrary to societal morality, it must not limit its inquiry merely to the defendant's appreciation that society, objectively speaking, condemned his actions. Rather, the jury should be instructed that it must also determine whether the defendant maintained a sincere belief that society would condone his actions under the circumstances as the defendant honestly perceived them. Finally, the trial court also should instruct the jury that, if it finds that the defendant had the substantial capacity to appreciate that his conduct both violated the criminal law and was contrary to society's moral standards, *even under the circumstances as he honestly perceives them*, then he should not be adjudged legally insane simply because, as a result of mental disease or defect, he elected to follow his own personal moral code.").

72. Perhaps the defendant recognizes that intentional homicide violates society's morals, but they feel compelled or an "irresistible impulse" to engage in the conduct. Although traditional moral wrongfulness

Jurisdictions have developed variations of the external morality model.[73] For example, the Connecticut Supreme Court endorses a version in which a defendant can be found legally insane if the defendant "substantially misperceived reality and harbored a delusional belief that society, *under the circumstances as the defendant honestly but mistakenly understood them,* would not have morally condemned his actions . . . ."[74] In the Connecticut version, a distorted perception of reality can extend not only to the factual circumstances of the conduct, but also to the defendant's understanding of what moral standards society actually holds.[75] In such a jurisdiction, the defendant can qualify for legal insanity on the basis of their delusional view of the facts and their potentially inaccurate view about whether society would approve of the conduct.[76]

In contrast, the Hawai'i Supreme Court has adopted a variation of the external morality model in which the jurisdiction asks "whether the defendant appreciated the wrongfulness of his conduct from the point of view of a reasonable person in the defendant's position under the circumstances as he believed them to be."[77] This variation of the external morality model starts with the defendant's perceived facts but shifts the focus away from the defendant's actual perception of morality and substitutes the perception of a reasonable person in the defendant's position. Here, although the facts of the incident may result from the delusion, it appears that only a reasonable person's perspective (and perhaps only reasonable mistakes) about morality and wrongfulness would be permitted to sustain an insanity claim.

Turning from state to federal treatment, in interpreting wrongfulness in the federal Insanity Defense Reform Act of 1984 (IDRA), the Seventh Circuit approved an instruction that asked "'whether the defendant appreciated that his

analyses do not support an insanity defense in these circumstances, jurisdictions that include an "irresistible impulse" provision could support an insanity defense here. "The volitional incapacity or irresistible-impulse test, which surfaced over two centuries ago (first in England, then in this country), asks whether a person was so lacking in volition due to a mental defect or illness that he could not have controlled his actions." Clark v. Arizona, 548 U.S. 735, 749 (2006) (footnotes omitted). Approximately 17 jurisdictions incorporate an irresistible impulse test (defendant cannot conform their conduct to the law). *See Kahler*, 140 S. Ct. at 1051–59 (Appendix): 1051, 1052–54, 1056–57. Others reject that test explicitly. *See, e.g.*, State v. Cordasco, 66 A.2d 27, 31 (1949) ("Insanity varying from this legal concept will not suffice as a defense and one who kills because of an irresistible impulse cannot seek an acquittal on that basis."); CAL. PENAL CODE § 28(b) (2003) ("As a matter of public policy there shall be no defense of diminished capacity, diminished responsibility, or irresistible impulse in a criminal action or juvenile adjudication hearing.").

73. *See e.g.*, *Wilson*, 700 A.2d at 643; State v. Uyesugi, 60 P.3d 843, 856 (Haw. 2002).

74. *Wilson*, 700 A.2d at 640, 643.

75. *Id.*; *see also* Yakush & Wolbransky, *supra* note 46, at 364.

76. However, by inquiring into the defendant's perception of society's approval of the conduct, the Connecticut version does maintain an external referent.

77. *Uyesugi*, 60 P.3d at 856 ("[W]e adopt a . . . subjective/objective rule [that] would determine whether the defendant appreciated the wrongfulness of his conduct from the point of view of a reasonable person in the defendant's position under the circumstances as he believed them to be.").

conduct was contrary to public morality.'"[78] This version also incorporates an external referent involving society's perspective on the wrongfulness of the act in question.

In contrast to an external morality model, an internal morality model focuses on whether the defendant believed that their conduct was morally justified under their own moral values. The Ninth Circuit in *United States v. Segna* described this approach to wrongfulness as being "'contrary to one's own conscience.'"[79] *Segna*, which predated the federal IDRA,[80] affirmed that, pursuant to the then existing doctrine, this type of internal morality model was the legally appropriate one.[81] Moreover, in 1999, fifteen years after the 1984 passage of the IDRA, a federal district court judge continued to find it appropriate to include analysis under an internal morality model[82] as one of three possible federal options, even as the court explicitly recognized that the internal morality evaluation was not likely to be the applicable one.[83]

Under an internal morality model, one might imagine, for example, a defendant who entertained a delusion, induced by a qualifying mental illness, that individuals with tattoos of a serpent were conspiring to control their life. In their delusional state, while having a dental cleaning, they see that the dental hygienist has a serpent tattoo and believe that the hygienist is installing a surveillance device in the defendant's teeth. As a result of the delusion, the defendant then grabs a cleaning instrument and assaults the hygienist. Assume that, under the defendant's personal moral code, violent prevention of the

78. United States v. Ewing, 494 F.3d 607, 612 (7th Cir. 2007); *see also* Yakush & Wolbransky, *supra* note 46, at 364. In a footnote, Yakush & Wolbransky opine that "it appears the federal system utilizes only an objective standard of morality. This is unlike California, which still allows the fact finder to assess both the defendant's subjective (whether he truly believed his act to be justified) and an objective standard (whether he understood that society would believe his act to be justified)." *Id.* at 364 n.13 (citation omitted).

79. United States v. Segna, 555 F.2d 226, 232 (9th Cir. 1977) ("Under this 'subjective' approach, the accused is not criminally responsible for his offending act if, because of mental disease or defect, he believes that he is morally justified in his conduct even though he may appreciate either that his act is criminal or that it is contrary to public morality.").

80. Insanity Defense, Insanity Defense Reform Act, 18 U.S.C. § 17(a) (1984) ("It is an affirmative defense to a prosecution under any Federal statute that, at the time of the commission of the acts constituting the offense, the defendant, as a result of a severe mental disease or defect, was unable to appreciate the nature and quality or the wrongfulness of his acts. Mental disease or defect does not otherwise constitute a defense.").

81. *Segna*, 555 F.2d at 232. The court in *Segna* was interpreting a version of the insanity standard modeled on a test developed by the American Law Institute (ALI) and relied on commentary from the ALI. *See Segna*, 555 F.2d at 232–33. The IDRA, however, did not adopt the ALI test. *See infra* note 84.

82. United States v. Danser, 110 F. Supp. 2d 807, 826–29 (S.D. Ind. 1999). The court does not use the terminology of this Article. Instead, the *Danser* court describes this model as "involving a purely subjective standard of morality[.]" *Id.* at 826, n.13.

83. *Danser*, 110 F. Supp. 2d at 827–29 (analyzing wrongfulness under the equivalent of external legality, external morality, and internal morality models and noting that the defendant "was able to appreciate the wrongfulness of his acts. This conclusion holds true regardless of which definition of 'wrongfulness' (as the term appears in 18 U.S.C. § 17) is used."). The *Danser* court noted that, because of the passage of the IDRA, "it seems only logical that Congress would not have intended the most expansive definition of 'wrongfulness' (under an already expansive insanity rule)—involving a purely subjective standard of morality—to now apply." *Id.* at 826, n.13.

implantation of a surveillance device was morally acceptable. Even if the defendant understood that society would view the conduct as illegal and criminal and would morally disapprove of the conduct, under an internal morality model, the defendant could be found insane.

Research suggests that courts have moved away from an internal morality model, and it is not clear that federal or state insanity doctrine currently officially adopts this approach.[84] Still, because it provides a logical possibility and due to its historical doctrinal recognition, this model offers a useful rubric for analysis and contrast to the external morality model.

B. LEGAL WRONG MODELS

Turning now from the two moral wrong to the two legal wrong models, we look first to an external legality approach. This would commonly involve ascertaining whether the defendant understood or could understand that the conduct violated society's law or was criminal. As a general rule, the defendant does not need to know the name of the crime or the specific law in question.[85] If, despite their mental illness, at the time of the act, the defendant understood or could understand that their conduct violated the law, this should generally result in rejection of the insanity defense under an external legality model.

---

84. *See, e.g.*, *Ewing*, 494 F.3d at 616–19 ("Ewing relies primarily on *United States v. Segna*, a decision by the Ninth Circuit, the only court to have adopted a subjective definition of wrongfulness like the one in Ewing's proposed instruction. There are a number of problems with reliance on *Segna*. First, the case predates the codification of the federal insanity defense and instead interprets wrongfulness as used in the *Model Penal Code*'s definition of legal insanity. That definition states: 'A person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity either to appreciate the criminality (wrongfulness) of his conduct or to conform his conduct to the requirements of law.' *Segna* relied principally on commentary from the American Legal Institute ('ALI') accompanying the formulation of this definition. But Congress did not adopt the *Model Penal Code*'s definition of insanity when it enacted the IDRA. Accordingly, neither the ALI commentary nor cases relying upon it are appropriate sources for interpretation of the statute. Moreover, although there is far from a robust body of case law on the issue, *Segna*'s subjective definition of wrongfulness-even in the context of § 4.01 of the *Model Penal Code* has been rejected by those courts to consider it since. . . . Accordingly, 'criminality' or 'contrary to law' is too narrow a definition of wrongfulness, and "subjective personal morality" is too broad. The second of the alternative definitions of wrongfulness-contrary to objective societal or public morality-best comports with the rules established in *M'Naghten's Case.*") (citations omitted) (footnotes omitted).

For an analysis by the Connecticut Supreme Court rejecting the internal moral wrongfulness approach, see State v. Wilson, 700 A.2d 633, 640 (Conn. 1997) ("We conclude that the defendant's efforts to define morality in purely personal terms are inconsistent with the Model Penal Code, judicial precedent, and the assumptions underlying our criminal law. . . . Although the rejection under the Model Penal Code of the personal standard is not beyond debate, we conclude that the drafters of § 4.01 did not intend that a defendant who appreciates *both* the illegality and the societal immorality of his actions be relieved of criminal responsibility due to his purely personal, albeit delusional, moral code. Moreover, the large majority of other jurisdictions that have considered the cognitive prong of the insanity defense has chosen a societal, rather than a personal, standard.") (citations omitted) (footnotes omitted).

85. *See, e.g.*, State v. Hamann, 285 N.W.2d 180, 183 (Iowa 1979) ("This is not to say, as has sometimes been suggested, that sanity would thereby be measured by legal knowledge. The test is not how much law a person claiming an insanity defense actually knows. The determination is to be made on the basis of a person's ability to understand it when something is prohibited by law.").

In contrast, under this model, the defendant could advance a legal justification or excuse that the defendant believed applied to their view of the facts and argue that, because of that justification, the defendant did not perceive the conduct as violating the law. If, with the facts as the defendant believed them to be, the conduct would not constitute a crime, an insanity defense could lie.[86]

A variation of Stephen's example of the homicide motivated by a need to save the human race, but in which the defendant does not understand the conduct to be legally wrong, could illustrate this situation.[87] Although in Stephen's era, under some circumstances, a need to save the human race could have provided a moral but not legal excuse[88] (because, inter alia, the legal defense of necessity did not generally extend to murder),[89] some jurisdictions today might permit the legal doctrine of necessity to justify (or mitigate) an intentional homicide.[90] Under an external legality model in such a jurisdiction, a defendant who believes their conduct is legally justified[91] and whose conduct, with the facts as the defendant understood them, meets the elements of that justification should be able to successfully invoke an insanity claim.[92] In this external legality model, the focus is on evaluation of whether the defendant's delusional state about the facts of the events, "if true, would lawfully justify" [93] the conduct. Under this formulation, the distortions that permit successful invocation of the defense generally extend only to the facts, not to the defendant's understanding of the jurisdiction's legal rules.

Within the context of legal wrongfulness, in contrast to the external legality model, one might posit the second model, an internal legality model, where the evaluation of the legal standard looks internally to the defendant's conception of society's rules.[94] This anticipates that the trier of fact would apply the

86. *See* Queen v. M'Naghten, (1843) 8 Eng. Rep. 718 (HL).

87. However, one would probably have to assume some additional facts, for example, the threat to extinguish the human race was clear and imminent, and there were no other legal alternatives to the defendant's conduct.

88. *See generally* Queen v. Dudley & Stephens, 14 Q.B.D. 273 DC (1884) (finding that necessity did not justify intentional homicide).

89. *See id.*

90. *See* Wood v. State, 271 S.W.3d 329, 334 (Tex. 2008) ("The jury was charged with finding Wood guilty of murder or, in the alternative, guilty of the lesser included offense of manslaughter, or not guilty. Murder is defined as intentionally or knowingly causing the death of another. Looking at the plain language of section 19.02, we do not perceive any legislative purpose indicating that the necessity defense is not available.") (footnotes omitted) (citation omitted).

91. Whether the defendant actually needs to believe that their conduct is legally justified, or whether the conduct just needs to meet the elements of the legal justification based on the defendant's perception of the facts, might be subject to debate or different variations of the external legality model.

92. For a functional application of an external legality model in a self-defense, rather than a necessity context, see, for example, People v. Leeds, 240 Cal. App. 4th 822, 824 (Cal. 2015) (holding that, in determining whether a defendant, who believed he was legally justified, could invoke a legal wrong insanity defense, the trier of fact could find "the defendant . . . legally insane if (1) he suffered from a delusional state and (2) his delusion, if true, would lawfully justify killing in self-defense.").

93. *Id.* at 824.

94. *See* Bloch & Gould, *supra* note 42, at 929. One could also posit a model in which the defendant looks only to their own legal constructs, ones that might be entirely divorced from laws of the outside world.

defendant's own perception of the jurisdiction's legal rules to the defendant's perception of the facts.[95] If a qualifying mental illness caused the defendant to fail or be unable to understand that they were violating the law, that could trigger qualification for an insanity claim here. This internal legality model provides for a defendant to misunderstand or misinterpret the law and still effectively invoke an insanity defense. Such mental illness-provoked misunderstanding could be a function of many influences,[96] and it would not require that the defendant have successfully fulfilled the actual elements of a legal justification or excuse.

Critics might contend that the above multi-model framework delves to a level of granularity and distinction that fails to acknowledge the confusion and breadth of distortions that mental illness can cause. But understanding right from wrong is contextual and nuanced and deserves elucidation to clarify distinctions both in perceptions by defendants and existing and potential relevant legal constructs.

In addition, critics might object because the analysis embeds the common assumption about free will underlying most criminal law—that free will exists and that human beings exercise choice in many contexts.[97] Neuroscientific inquiry is raising questions about the existence and applicability of free will.[98] Consequently, it might be that criminal law doctrine fails to reflect the reality of how our brains function. Such a failure would suggest that our current doctrinal

---

Commentators might disagree about whether such a model does or could exist. *See, e.g.*, Yakush & Wolbransky, *supra* note 46, at 359–60 ("[W]ithin any given jurisdiction, there is only one set of laws. Thus, in regard to the legal aspect of wrongfulness, a citizen of California is judged by whether or not he had the capacity to know that the act violated the laws of that state. . . . [T]here is only one perspective of what is illegal and legal.").

95. *See* Bloch & Gould, *supra* note 42, at 929.

96. For a discussion of some of the influences that might produce such a misperception and/or that might provide the defendant a personal perception that the basic tenets of a legal justification were met, see e.g., Johnston & Leahey, *supra* note 10, at 1811–12, 1820–36.

97. Morissette v. United States, 342 U.S. 246, 250 (1952) (explaining that "[t]he contention that an injury can amount to a crime only when inflicted by intention is no provincial or transient notion. It is as universal and persistent in mature systems of law as belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil.").

98. *See, e.g.*, Robert M. Sapolsky, *Neuroscience and the Law*, 15 U. ST. THOMAS J. L. & PUB. POL'Y, 138, 138 (2021) ("I believe there is no free will whatsoever, and that is going to have to utterly transform how we think about every aspect of our society. From how we judge harshly, to how we praise, and everything in between."). *See generally* Patrick Haggard, *Decision Time for Free Will*, 69 NEURON 404 (2011); Peter S. Davies & Peter A. Alces, *Neuroscience Changes More than You Can Think*, 2017 U. ILL. J.L. TECH. & POL'Y 141 (2017) (reviewing OWEN D. JONES, JEFFREY D. SCHALL & FRANCIS X. SHEN, LAW & NEUROSCIENCE (2014). *But see e.g.*, Stephen J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology*, 25 BEHAV. SCI. L. 203, 203 (2007) ("Solving the free will problem would have profound implications for responsibility doctrines and practices, but, at present, the problem plays no proper role in forensic practice or theory because this ability or its lack is not a criterion of any civil or criminal law doctrine."); Stephen J. Morse, *Internal and External Challenges to Culpability*, 53 ARIZ. ST. L.J. 617, 618 (2021) ("I address the newer, broader challenges to personhood, agency and responsibility that are fueled by alleged advances in behavioral neuroscience and genetics. Some of these are quite radical. They may even turn out to be correct, but at present, there is no conceptual or empirical reason to believe that they are true. Moreover, there is certainly insufficient reason to jettison notions of criminal responsibility that have been developing for centuries and to adopt instead the proposed, radical conception of justice.").

approach to insanity, among other criminal law principles, might need reevaluation.[99]

In the interim, the models offer a range of variations in the understanding of moral and legal wrongfulness, a number of which reflect different doctrinal approaches to legal insanity in the United States today.

II. IMPLICATIONS OF THE DOG, WOLVES, AND GOD EXAMPLES

With this conceptual and doctrinal backdrop, we turn now to parse in some detail the dissent's and majority's discussions that suggest a dog's command to kill could enable a successful insanity claim. The dog-as-commander example implies two critical assumptions. The first is important but relatively uncontroversial. If one should be excused under legal insanity when one perceives a dog as ordering one to commit homicide,[100] this suggests, as described earlier,[101] that the trier of fact should take the facts as the defendant believed them to be when evaluating an insanity claim. This reflects and endorses a fundamental principle of legal insanity and delusions as articulated in *M'Naghten*: the jury looks through the defendant's lens and delusion, no matter how distorted, to ascertain the facts of the events to which the law will be applied.[102]

The second assumption is that a delusion unmoored to societal perceptions of wrongfulness (either moral or legal) necessarily warrants a finding of legal insanity. That assumption and its problematic implications drive this Article. If we declare that someone implementing a dog's command to commit homicide, even someone claiming moral correctness and acting in response to a qualifying mental illness, is necessarily legally insane, what does that mean in terms of the moral capacity prong of *M'Naghten* or its progeny?

Three concerns emerge from the approach. First, the dissent implies that obeying the mental-illness generated command of a dog to kill equates to an inability to discern right from wrong.[103] In his example where the dog commands the killing, Justice Breyer argues that "the defendant[] differ[s] from ordinary persons in ways that would lead most of us to say that they should not be held morally responsible for their acts."[104] But the only significant difference the dissent has enumerated is that the defendant suffers from the command hallucination caused by a mental illness.[105] From a conceptual perspective,

99. *See, e.g.*, Gardar Árnason, *Neuroscience, Free Will and Moral Responsibility*, 15 J. HUMANS. & SOC. SCI. 147, 154 (2011). Árnason opines that "[i]f the view that free will is an illusion becomes widely accepted . . . the legal system would have to be drastically revised. Justice and desert cannot play any part in punishment . . . " *Id.* at 147. Árnason argues, however, "that neuroscience has not revealed free will to be an illusion and that it is not likely ever to do so." *Id* at 154.

100. Assuming the trier of fact finds that the defendant believed they actually heard this command.

101. *See supra* note 62 and accompanying text.

102. *Id.*

103. Kahler v. Kansas, 140 S. Ct. 1021, 1038 (Breyer, J., dissenting).

104. *Id.* at 1048.

105. *Id.* at 1038–50.

determining wrongfulness, however, depends upon more than that. It depends on which model or variation the jurisdiction applies to evaluate insanity claims. A blanket conclusion that the defendant qualifies as legally insane because a dog commanded the defendant to kill, without an investigation of the applicable model, is premature and potentially inaccurate. Second, and, in addition, the defendant's motivation for acting in response to the perceived command can be determinative of an insanity claim. The dissent supplies no rationale for the defendant's acquiescing to the command (other than perhaps the mental illness itself and that the example "falls within [*M'Naghten's*] second prong"),[106] which also suggests that Justice Breyer's conclusion is premature. Third, if the majority's exposition of the example,[107] is somehow understood as endorsing an internal morality model, that embrace warrants explicit consideration. The majority opinion refers to "those who know exactly what they are doing (including that it is against the law) but believe it morally justified—because, say, it is commanded by God (or in the dissent's case, a dog)."[108] In particular, a defendant believing their conduct is morally justified when acting on the command of a dog leaves open the question of whether their belief takes into account an external morality or stems from a personal internal moral code.

In *Kahler*, the Court recognized that jurisdictions have adopted myriad versions of the early *M'Naghten* test, with some closer to the original, and others rejecting or substantially rewriting portions of it.[109] As the *Kahler* dissent and Appendix make clear, however, the vast majority of U.S. jurisdictions continue to include a wrongfulness prong of some variety.[110] But the fact that there is a wrongfulness prong does not answer the question of whether the defendant who obeys the dog will qualify as insane pursuant to the specific wrongfulness model adopted. Consequently, without contextualizing which variation of the models applies and how, readers are left without adequate guidance.

Beyond the importance of knowing the specifics of the jurisdiction's adopted model, Justice Breyer's underspecified example in the dissent lacks any discussion of motivational incentives beyond the fact that the conduct was commanded by a dog.[111] Even if the individual's motivation derives only from

106. *Id.* at 1038–1050.

107. *Id.* at 1036 (majority opinion).

108. *Id.*

109. *See id.* at 1025.

110. *Id.* at 1051–59 (Appendix).

111. *Id.* at 1038–50 (Breyer, J., dissenting). Justice Breyer does suggest that the illness-generated command delusion means that the accused does not understand the wrongfulness of the conduct, but this seems to be a conclusion, rather than a motivation for the conduct as that conduct is described in the dissent's analysis. *Id.* In Justice Breyer's dissent from the denial of certiorari in *Delling v. Idaho*, Justice Breyer cites an amicus brief, which explains that mental illness can cause an individual to "'wrongly believe the act is justified.'" Delling v. Idaho, 568 U.S. 1038 (2012) (Breyer, J., dissenting in denial of certiorari) (citing Brief Amici Curiae of 52 Criminal Law and Mental Health Law Professors at 10, *Delling*, 568 U.S. 1038 (No. 11-1515)). However, Justice Breyer does not describe the illness in the dog-as-commander example beyond noting that it is "a severe mental illness." *Id.* at 1038.

a delusional view of reality,[112] rarely will an individual act without some motivation. Noting that the defendant's acts were prompted by the dog's command does provide an initial level of motivation, but the command by the dog to commit homicide tells us little about why the defendant obeyed. If the defendant obeyed because the defendant has enormous respect for the intelligence of dogs, that might produce a different result than if, similar to Stephen's excerpted example, the defendant claimed that the dog was an oracle who could predict the future and, in that future, all of the human race would perish immediately if the defendant did not commit the killing. Arguably, under the respect for canine intelligence motive, even with the facts as the defendant understood them, the delusion invokes no traditional defense or excuse, nor does it effectively suggest that the defendant failed to or was unable to understand the wrongfulness (moral or legal) of the act. In an external legality or morality model, that presentation would not necessarily lead to the dissent's conclusion that an insanity claim would prevail. As previous scholars have opined more generally, "a defendant who experienced command auditory hallucinations that ordered him to commit an illegal/immoral act, while retaining the cognitive reasoning necessary to know it was nonetheless wrong, would have no case for insanity."[113] In contrast, believing that the conduct was morally or legally right, because it would preserve the human race from imminent destruction, might engage a moral or legal defense of others or necessity[114] claim. There, an insanity verdict—a defendant's failure to understand that the conduct was wrong under some moral or legal societal standard—arguably possesses logical force in an external model.

Motivations can be determinative of the insanity claim. Although Justice Breyer emphasizes the importance of a motivational inquiry,[115] beyond acknowledging that the defendant was responding to a dog's command, the example fails to respect the key role played by asking "why" the defendant perpetrated the act. This Article suggests that the lack of adequate inquiry about motivational rationale undermines Justice Breyer's definitive conclusion that legal insanity should excuse the defendant in the dog-as-commander example.

112. *Kahler*, 140 S. Ct. at 1048 (Breyer, J., dissenting). Justice Breyer notes that "mental illness . . . affects [] *motivations* for forming [] intent." *Id.* (citation omitted). He writes, "[f]or example, the American Psychiatric Association tells us that individuals suffering from mental illness may experience delusions—erroneous perceptions of the outside world held with strong conviction. They may believe, incorrectly, that others are threatening them harm (persecutory delusions), that God has commanded them to engage in certain conduct (religious delusions), or that they or others are condemned to a life of suffering (depressive delusions)." *Id.* (citation omitted).

113. Yakush & Wolbransky, *supra* note 46, at 366. Yakush & Wolbransky do provide a separate analysis of commands from God or "a higher power." *Id.* at 370–71.

114. The necessity or lesser-of-two-evils option, however, would presumably only be available in a jurisdiction that allowed such a defense in homicide cases and then only under the requirements of that jurisdiction's approach to insanity.

115. *Kahler*, 140 S. Ct. at 1048 (Breyer, J., dissenting).

Let us consider more specifically how two of the models outlined earlier might apply to the dog-as-commander example to demonstrate the potential inaccuracy of the legal insanity label that the dissent assigns and the Court implies. The analysis applies an externally-focused morality framework and then an internally-focused morality framework.[116]

In an external morality evaluation, we generally assess whether, relying on the defendant's perception of the facts, the defendant understood that killing a human being would violate society's code of appropriate conduct or society would find that conduct morally unacceptable.[117] Even if the defendant perceived that a dog ordered the defendant to kill someone and had no personal moral qualms about the killing, the defendant might, for example, still understand that society would condemn such a killing.[118] The defendant might have obeyed because they respected canine intelligence or wanted to demonstrate the type of reciprocal loyalty to canines that dogs often display toward their owners. Alternatively, the defendant may have obeyed for innumerable other reasons, like an inability to conform their conduct to the law,[119] which would not necessarily defeat or undermine the defendant's ability to understand that homicide under such circumstances would violate society's moral codes or norms. In these scenarios, even if the defendant personally believed that the killing was morally justified, the defendant would be unlikely to meet the doctrinal prerequisites for legal insanity under an external morality approach.

---

116. This can be a multifaceted analysis when the defendant claims to have "special knowledge" as might be the case if the defendant claimed that the dog was an oracle who could predict the future and, in that future, without the death of the homicide victim, all of the human race would perish immediately.

117. *See e.g.*, *supra* notes 74–78 and accompanying text. Yakush & Wolbransky suggest that an additional inquiry about the defendant's personal belief that the conduct was justified may also be involved in California. Yakush & Wolbransky, *supra* note 46 at 364 n.13.

118. Professor Fredrick Vars offers another perspective on Justice Breyer's failure to analyze the perception of wrongfulness in the dog example. Fredrick E. Vars, *Of Death and Delusion: What Survives* Kahler v. Kansas*?*, 169 U. PA. L. REV. ONLINE 90, 93 (2020) (footnote omitted). He suggests that "[a]s [this] hypothetical reveals, the dissent was not wed to the moral incapacity test: one could feel compelled to follow a dog's order even while recognizing that doing so is wrong." *Id. But see* Morse, *Internal and External Challenges*, *supra* note 98, at 645 ("One of the defects of the *Kahler* dissent, in my opinion, is that it suggested that some form of 'moral understanding' test was required.").

119. This test is generally known as the "irresistible-impulse" test. *Kahler*, 140 S. Ct. at 1028. Trial testimony from the defense's forensic psychiatrist in the *Kahler* case raised this possibility. The psychiatrist opined that Kahler's "'capacity to manage his own behavior had been severely degraded so that he couldn't refrain from doing what he did.'" State v. Kahler, 410 P.3d 105, 114 (Kan. 2018). Professor Joshua Dressler has argued: "James Kahler likely was *not* insane under the *M'Naghten* test. The closest the defense expert came to making a case for insanity was when he reported that Kahler 'felt compelled' to kill. That sounds like a volitional incapacity claim, not a cognitive one. And, since Kahler purposely spared his son—who wanted to spend more time with his father—this fact makes even a volitional argument suspect." Joshua Dressler, Kahler v. Kansas*: Ask the Wrong Question, You Get the Wrong Answer*, 18 OHIO ST. J. CRIM. L. 409, 413 (2020) (footnotes omitted). Other commentators also note that expert testimony at Kahler's trial suggested that Kahler might "have become dissociated from reality at the time of the crime." Morse, *Internal and External Challenges*, *supra* note 98, at 643.

The analysis thus far suggests that the dissent's conclusion, that the dog-as-commander example as given would result in an insanity verdict, is likely inaccurate under an external morality model. This is because experiencing an illness-generated command from a dog to kill does not necessarily mean that the defendant believed that society would approve of the conduct nor does the dissent offer a motivation or rationale that would provide society, in the form of the trier of fact, a moral justification, beyond the mental illness itself, for approving the conduct.

With respect to the *Kahler* majority opinion, the majority does expressly recognize that the defendant's perception of morality matters.[120] However, the majority's assertion that the defendant believed their conduct to be "morally justified"[121] in the dog or God examples does not resolve the concern about whether that belief involved reference to societal norms or to the defendant's personal moral values. While many readers might fairly assume an external morality model, others might not. In light of the history of the application of the internal morality model in at least one federal circuit court,[122] this lack of clarity calls for further elucidation by the majority. In addition, in some applications, even if the defendant believed that society would not condemn their actions, that is not the determinative test under that external morality model.[123] Consequently, even if one interprets the majority's description as referring not to the defendant's personal morality but to the defendant believing that society would find their conduct morally justified, if the defendant's conduct of obeying the dog actually violated societal norms from the perspective of a reasonable person in the defendant's circumstances, the defendant's belief would be unlikely to excuse the defendant's conduct in a jurisdiction with such a definition. Both of these limitations, the lack of specification of societal as opposed to personal morality and the existence of a jurisdictional approach to external morality that would not necessarily excuse the conduct, call for greater clarity in the treatment of wrongfulness by the *Kahler* majority.

Beyond the questions of the rationale, the defendant's awareness of wrongfulness, internal versus external morality, and which definition of morality controls in an external morality model, consider more specifically the question of the identity of the commander. Imagine that, instead of perceiving a dog as commanding the killing, the defendant perceived that a human stranger ordered them to commit the murder (and that perception was due to a hallucination provoked by a qualifying mental illness). Would the defendant necessarily qualify for an insanity defense? Among other considerations, the answer should require a determination of the wrongfulness model applicable in the jurisdiction, a motivational analysis, and, depending on the jurisdiction, a careful assessment

120. *Kahler*, 140 S. Ct. at 1036.

121. *Id.*

122. *See supra* notes 79–83 and accompanying text.

123. *See e.g.*, State v. Uyesugi, 60 P.3d 843, 856 (Haw. 2002).

of what the defendant understood at the time about the wrongfulness of the conduct. The key difference between the dog and stranger examples is that society would probably assume mental illness existed when a defendant claimed to have acted in response to a command from a dog,[124] as most of us do not perceive dogs as ordering us to behave, whereas that assumption does not apply in the case of commands by other humans. Using a dog as the commander makes visible the likelihood of mental illness, but does not ordain the conclusion of legal insanity. Thus, contrary to Justice Breyer's assertion in the dissent, on this analysis too, insanity doctrine in jurisdictions that apply an external view of moral wrongfulness would not necessarily support an insanity verdict here.

In contrast to a command from a dog, if the defendant perceived the command as emanating from God, the analysis would likely change. As courts and commentators have noted, a perceived command from God can often be understood as invoking a higher law.[125] A defendant obeying such a command might well believe that their conduct in response to the command would not violate society's moral codes or norms.[126] Moreover, society, through the triers of fact, might readily perceive obeying a command that the defendant understood came from God as within the norms of society. Thus, under an external morality approach, they could conclude that the defendant failed to perceive the conduct as wrong and condemned by society's morals, and they too might decide that, under the circumstances, responding to such a command did not violate society's norms. Under this scenario, the verdict might well be in conformity with the assertion of legal insanity in *Kahler*. The actor who does the commanding can matter in the insanity context. Not only does the dissent's choice of a dog as the commander raise questions about whether an insanity verdict is likely in an external morality model, but, in addition, the majority's failure to explicitly distinguish between a command from a dog and one from God renders the opinion subject to criticism here with respect to that model.[127]

Returning to the command by a dog, but next superimposing an internal morality model, this model asks whether the defendant was able to understand

124. This is consistent with the approach by Justice Breyer. *See Kahler*, 140 S. Ct. at 1048 (Breyer, J., dissenting).

125. *See, e.g.*, People v. Schmidt, 110 N.E. 945 (N.Y. 1915) ("Obedience to the law is itself a moral duty. If, however, there is an insane delusion that God has appeared to the defendant and ordained the commission of a crime, we think it cannot be said of the offender that he knows the act to be wrong."); Brief of American Psychiatric Ass'n. & American Academy of Psychiatry and the Law as Amici Curiae in Support of Petitioner at 16, Delling v. Idaho, 568 U.S. 1038 (2012) (No. 11-1515), ("And, religious delusions can be manifested as a conviction that one must carry out certain acts – even if they are against the law – because they are commanded by God.") (citation omitted); State v. Wilson, 700 A.2d 633, 641 (Conn. 1997) ("[W]e are hard pressed to envision an individual who, because of mental disease or defect, truly believes that a divine power has authorized his actions, but, at the same time, also truly believes that such actions are immoral. An individual laboring under a delusion that causes him to believe in the divine approbation of his conduct is an individual who, in all practicality, is unlikely to be able fully to appreciate the wrongfulness of that conduct.").

126. *Cf. Wilson*, 700 A.2d 633, 641.

127. The majority opinion had a focused discussion about commands perceived as coming from God, which, in a jurisdiction that had a moral capacity option, might excuse a defendant. *See Kahler*, 140 S. Ct. at 1026.

that their actions were wrong according to their own internal world view, not whether their behavior violated society's moral codes. This model might upend current state and federal interpretations of the wrongfulness prong.[128] However, assuming it were the relevant model, although Justice Breyer asserts that the dog-as-commander example falls within the *M'Naghten* prong on wrongfulness, the *Kahler* dissent does not offer an explicit analysis of the defendant's perception of wrongfulness in its use of the dog analogy.[129] Consequently, it is not clear that application of this model in the dissent's dog analogy would produce a legal insanity result here either.

In the *Kahler* dissent, for an internal morality model, where the defendant's perception of the morality or wrongfulness of, and rationale for, obeying the command were not detailed, to find insanity, one presumably must accept that following a command from a dog to kill a human does not register on the defendant's moral compass as wrong. Without more, this seems a questionable assumption. Even if one dwells within the defendant's internal universe of acceptable behavior, and accepts as true that dogs' commands must be obeyed, the defendant might still find their own conduct reprehensible. Maybe the defendant understands the nature and quality of their act and understands their conduct as wrong (legally and morally) but cannot help themselves from committing it. As only a minority of jurisdictions incorporate an irresistible impulse prong into their insanity tests,[130] and that test involves a different inquiry than the one in the internal morality model described here, the defendant might be found sane under even an internal morality wrongfulness rubric. Overall, the dissent's assertion that a dog's homicidal command to the defendant would necessarily excuse the defendant under the traditional doctrine of legal insanity is not borne out by careful analysis of current external morality doctrinal approaches or perhaps even an internal morality approach.

To be clear, Justice Breyer's dissent focuses on the importance of incorporating an analysis of wrongfulness in sanity evaluations to excuse a defendant and decries the majority's finding that Kansas can constitutionally permit the elimination of that traditional analysis.[131] Despite this, however, the simplified dog example lacks critical analysis of what wrongfulness means.

---

128. United States. v. Ewing, 494 F.3d 607, 610 (7th Cir. 2007) ("*M'Naghten's Case* and American case law applying it establish that a defendant's ability to appreciate right and wrong has consistently been determined by reference to societal, not personal, standards of morality.").

129. *Kahler*, 140 S. Ct. at 1025, 1038 (Breyer, J., dissenting). Justice Breyer does write that "[i]t is no defense simply to claim that one's criminal conduct was morally right." *Id.* at 1047. However, this does not appear to refer to definitions of wrongfulness in insanity. Rather, it appears in the context of discussing the relationship more generally of moral law to the criminal law, as follows: "The criminal law does not adopt, nor does it perfectly track, moral law. It is no defense simply to claim that one's criminal conduct was morally right. But the criminal law nonetheless tries in various ways to prevent the distance between criminal law and morality from becoming too great." *Id.* Included among the ways that Justice Breyer explains that distance is mitigated are defenses, like insanity and duress, prosecutorial discretion to decline prosecution, and executive clemency. *Id.*

130. For a brief discussion of the irresistible impulse test, see *supra* note 72.

131. *Kahler*, 140 S. Ct. at 1038–50 (Breyer, J., dissenting).

Before concluding, we should talk about supernatural wolves. In the 2012 dissent to a denial of certiorari in *Delling v. Idaho*,[132] Justice Breyer furnished a different but related "very much simplified example"[133] of legal insanity. Justice Breyer posited: if "the defendant, due to insanity, believes that a wolf, a supernatural figure, has ordered him to kill the victim [,]"[134] that would furnish a legal insanity defense.[135] Two significant features of this example and the explanation that Justice Breyer provides distinguish it from the dog example in *Kahler*.

First, Justice Breyer provides an important insight into the defendant's perspective about the wolf. Unlike the dog, the wolf is not from this world. A supernatural entity is defined as one "of or relating to an order of existence beyond the visible observable universe *especially*: of or relating to God or a god, demigod, spirit, or devil."[136] In other words, the supernatural wolf figure could be understood as invoking some of the deference and powers associated with powerful entities or deities of religion whose commands are often perceived as morally compelling. Thus, the supernatural wolf might be seen as more akin to a deity in terms of the impact on the insanity analysis than to a dog that lacks any attribution of extraordinary power.

The second feature that the *Delling* dissent offers is a hypothesis to explain the defendant's belief that the act was not wrongful. Justice Breyer indicates that "*amici* tell us that those seriously mentally ill individuals often possess the kind of mental disease . . . [where] they know that the victim is a human being, but due to mental illness, such as a paranoid delusion, they wrongly believe the act is justified."[137] Providing that rationale represents greater engagement with essential dimensions of wrongfulness in legal insanity than does the treatment of the dog-as-commander example by the dissent in *Kahler*.

Still, even with reference to the supernatural quality of the wolf and a recognition that mental illness can cause an inaccurate perception of wrongfulness, the wolf example should not automatically warrant an insanity defense. As with the dog example, the wolf example lacks important information, including recognition that multiple wrongfulness models exist and that their differences matter. For example, in a jurisdiction that applies an external legality model, a defendant who believes their act was morally justified might still understand that an order to commit murder, even from a supernatural

132. Delling v. Idaho, 568 U.S. 1038, 133 S. Ct. 504 (2012) (Breyer, J., dissenting in denial of certiorari).

133. *Id.* at 505.

134. *Id.*

135. *Id.*

136. Definition of "supernatural," MERRIAM-WEBSTER DICTIONARY, https://www.merriam-webster.com/dictionary/supernatural (last visited Apr. 15, 2022).

137. *Delling*, 568 U.S. 1038, 133 S. Ct. 504 (Breyer, J., dissenting in denial of certiorari). In the *Kahler* dissent, Justice Breyer also refers to information from the Brief for American Psychiatric Association et al. as *Amici Curiae*, but the *Kahler* reference lacks a focus on the question of the defendant's perception of moral justification. Kahler v. Kansas, 140 S. Ct. 1021, 1048 (2020) (Breyer, J., dissenting).

wolf, violates society's laws.[138] In that jurisdiction, with this understanding, contrary to Justice Breyer's conclusion, the defendant would be unlikely to qualify as insane.[139]

CONCLUSION

Why does it matter that the Supreme Court's engagement with the simplified dog-as-commander example in *Kahler* lacks analytical depth and precision? Supreme Court language and examples provide grounding for lower court interpretation and real-world implementation, even when they are not the holding. Consequently, unearthing limitations of the Court's and dissent's analyses can deter reliance on problematic implications of the dog-as-commander example. Wrongfulness analyses benefit from careful parsing and application of the jurisdiction's doctrinal model, as well as a detailed focus on why the defendant engaged in the conduct at issue. The challenges and opacity that *Kahler*'s invocation of the dog-as-commander engenders call for further efforts to clarify the intended scope and implications of wrongfulness in legal insanity.

138. Because an external legality model still anticipates a wrongfulness or blameworthiness evaluation, and because Justice Breyer argues at some length that, although logically different, "there is no indication that [the distinction between legal and moral wrong] makes a meaningful difference in practice[,]" it seems fair to evaluate the dog-as-commander example under this model. *Id.* at 1046.

139. Alternatively, even if the defendant believed that society would not condemn their act, they might find themselves in a jurisdiction that anticipated evaluation of wrongfulness from the perspective of a reasonable person in the defendant's circumstances. *See supra* note 77 and accompanying text. That perspective might produce a different result than the insanity determination suggested by Justice Breyer in the dog-as-commander example in the dissent in *Kahler*.