THE HEAT OF PASSION AND BLAMEWORTHY REASONS TO BE ANGRY

Jonathan Witmer-Rich*

Introduction

This article seeks to resolve a longstanding conceptual puzzle plaguing the "heat of passion" doctrine—how courts should determine which features, beliefs, or characteristics of a defendant are properly relevant to assessing whether the defendant was sufficiently provoked, and which of those features should be disregarded. This article argues that provocation is not adequate if the reason the defendant became extremely angry is due to some blameworthy belief or attribute of the defendant. A belief is blameworthy if it contradicts the fundamental values of the political community. The blameworthiness principle distinguishes those aspects of the defendant that cannot form a basis to argue the defendant was reasonably provoked from those aspects that can properly form the basis of a provocation claim.

Part I introduces the "heat of passion" doctrine and briefly explains both the conceptual problem and this article's solution. Part II more carefully examines the nature of the "heat of passion" doctrine, focusing in particular on the second component—whether the provocation was "reasonable" or "sufficient." Part III frames the problem—which characteristics of a defendant should be considered when assessing the reasonableness of provocation—by describing four example cases that elicit differing intuitions among many commentators. Part IV considers standards articulated by other commentators seeking to resolve this problem. Using the examples given in Part III, Part IV also explains why each of these standards fails to properly differentiate the cases. Part V then explains the correct standard—that the sufficiency of the provocation must be based on whether the defendant's reason for becoming extremely angry is itself blameworthy. Part V also explains how this standard correctly distinguishes among the examples given, and is grounded on the fundamental principle of culpability underlying criminal punishment. Part VI turns to several recurring types of provocation claims: claims by men who kill their intimate partners, and so-called "gay panic" and "trans

---

* Joseph C. Hostetler—BakerHostetler Professor of Law, Cleveland-Marshall College of Law, Cleveland State University; J.D., University of Michigan School of Law; B.A., Goshen College. Thanks to those who graciously commented on drafts of this article: Eric Blumenson, Michal Buchhandler-Raphael, Michael T. Cahill, Joshua Dressler, Avlana K. Eisenberg, Stephen P. Garvey, Steven R. Morrison, and Peter Westen; participants in the CrimFest 2015 Conference, in particular Michelle Anderson, Anthony Dillof, Cynthia Godsoe, Benjamin Levin, Michael J.Z. Mannheimer, Alice Ristroph, and Laurent Sacharoff; and participants in presentations to the faculty at the DePaul University College of Law and the Cleveland-Marshall College of Law.

panic" cases. Part VI also explains how these types of cases likewise involve the same conceptual issue at play in cases involving minority cultural or religious beliefs and how the "blameworthiness" principle provides the correct standard for differentiating the few valid provocation claims in these cases from the many invalid claims.

I. THE PROVOCATION OR "HEAT OF PASSION" DOCTRINE

In the criminal law, the "heat of passion" doctrine is a partial defense that reduces an offense that would otherwise constitute murder to the lesser offense of manslaughter.[1] Also called the "provocation" doctrine, this partial defense requires the defendant to show the following: (1) "[t]he actor must have killed while in the heat of passion"; and (2) "[t]he heat of passion must have been brought about by adequate provocation."[2]

Courts and commentators regularly state that the heat of passion doctrine contains both subjective and objective components.[3] The "heat of passion" element is subjective—the defendant must show that at the time of the killing, he was in fact in a state of passion or extreme anger.[4] The "sufficient provocation" element is objective, in that it depends on evaluating the defendant's extreme anger against some objective normative standard. (The nature of that normative evaluation is discussed below.) If a defendant is extremely angry but lacks sufficient provocation, the partial defense is not available. If a defendant is extremely angry and there is sufficient provocation, the partial defense is satisfied, and the defendant is guilty of manslaughter rather than murder.

---

1. *See* JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW § 31.07, at 530–41 (7th ed. 2015); WAYNE R. LAFAVE, CRIMINAL LAW § 15.2, at 1043–44 (6th ed. 2017).

2. Mitchell N. Berman & Ian P. Farrell, *Provocation Manslaughter as Partial Justification and Partial Excuse*, 52 WM. & MARY L. REV. 1027, 1040 (2011).

3. Joshua Dressler, *Rethinking Heat of Passion: A Defense in Search of a Rationale*, 73 J. CRIM. L. & CRIMINOLOGY 421, 427 (1982) ("The provocation defense came to include both subjective and objective elements."); *see* California v. Oropeza, 59 Cal. Rptr. 3d 653, 660 (Ct. App. 2007) ("Voluntary manslaughter has both a subjective and an objective requirement. The defendant must kill while actually in the heat of passion. That heat of passion, however, must be aroused by sufficient provocation judged objectively."); State v. Shane, 590 N.E.2d 272, 276–77 (Ohio 1992); State v. Williford, 307 N.W.2d 277, 283 (Wis. 1981).

4. Jurisdictions differ in how they define what sort of state the defendant must be in. Some say "heat of passion." *See* ALASKA STAT. § 11.41.115(a) (2016); MINN. STAT. ANN. § 609.20(1) (West 2018). Some say "extreme emotional disturbance." *See* ARK. CODE ANN. § 5-10-104(a)(1)(A) (2016); KY. REV. STAT. ANN. § 507.020(1)(a) (West 2017); MODEL PENAL CODE § 210.3 (AM. LAW INST. 1985). Some use formulations like "sudden passion or . . . sudden fit of rage," OHIO REV. CODE ANN. § 2903.03(A) (West 2017), or "extreme anger or extreme fear," ME. STAT. tit. 17-A, § 203(1)(B) (2017). This article does not take a position on which of these emotional states is correct.

A. The Problem: How to Individualize the Reasonable Person

One of the questions a fact-finder must answer, in considering this partial defense, is whether the provocation was "sufficient"[5] or "reasonable"[6] or "adequate."[7] Under this standard, there will be some cases in which a defendant was indeed provoked into extreme anger, but it was not reasonable or appropriate for him to have become so angry. Or perhaps the standard suggests that the cause of his extreme passion was not adequate or sufficient to justify his violent conduct.

How, then, should fact-finders determine whether there was "reasonable provocation" in a given case? One conventional first step, relying on the "reasonable person" who appears throughout the law, is to ask, "would a reasonable person have been provoked into a heat of passion?"[8]

A recurring problem with this formulation is the difficulty of how much to "individualize" the reasonable person—how to determine which characteristics of the defendant (physical traits, emotional dispositions, past experiences, beliefs, etc.) should be imported into this "reasonable person."[9] Obviously, the "reasonable person" must be a person who has just experienced what the defendant experienced. Asking whether a "reasonable person" would have become extremely provoked at the sight of his child being sexually assaulted presupposes that the reasonable person is a parent who has a child, and the reasonable person has just witnessed that child being sexually assaulted.

---

5. *See* MISS. CODE ANN. § 97–3–17 (2017) ("sufficient provocation"); State v. Wharton, 672 S.E.2d 786, 788 (S.C. 2009) ("sufficient legal provocation").

6. *See* N.J. STAT. ANN. § 2C:11-4(b)(2) (West 2017) ("reasonable provocation").

7. *See* ARIZ. REV. STAT. ANN. § 13-1103(A)(2) (2017) ("adequate provocation").

8. *See* Commonwealth v. Vatcher, 781 N.E.2d 1277, 1281 (Mass. 2003) ("the jury must be able to infer that a reasonable person would have become sufficiently provoked. . . ."); State v. Smith, 858 N.E.2d 1222, 1238 (Ohio Ct. App. 2006) (considering whether the victim's conduct "had the potential to provoke an objectively reasonable person into a sudden fit of passion or rage"); McClung v. Commonwealth, 212 S.E.2d 290, 292 (Va. 1975) ("[T]he jury could have concluded that his conduct under these circumstances was sufficient to provoke a reasonable person to fear or rage, or both."); Dressler, *supra* note 3, at 427; Cynthia Lee, *The Gay Panic Defense*, 42 U.C. DAVIS L. REV. 471, 500 (2008) (noting that jurisdictions commonly ask "if the reasonable person in the defendant's shoes would have been provoked into a heat of passion"); Peter Westen, *Individualizing the Reasonable Person in Criminal Law*, 2 CRIM. L. & PHIL. 137, 139 (2008).

9. CYNTHIA LEE, MURDER AND THE REASONABLE MAN: PASSION AND FEAR IN THE CRIMINAL COURTROOM 204–12 (2003); Guy Ben-David, *The Integration of Cultural Arguments in the Provocation Doctrine in Rulings of the Israeli Supreme Court*, 2 CREIGHTON INT'L & COMP. L.J. 1, 2 (2011) (highlighting "the question of which of the defendant's characteristics, if any, should be taken into account to estimate the intensity of the provocation. It may be that a person who is a part of a minority culture will be provoked by words or actions that would not subjectively provoke someone who is not a part of that same culture"); Joshua Dressler, *Why Keep the Provocation Defense?: Some Reflections on A Difficult Subject*, 86 MINN. L. REV. 959, 986–87 (2002) (noting "an exceptionally troubling issue. . . . To what extent are a defendant's subjective characteristics relevant in applying the objective 'reasonable person' standard" in provocation cases); Orit Kamir, *Responsibility Determination as a Smokescreen: Provocation and the Reasonable Person in the Israeli Supreme Court*, 2 OHIO ST. J. CRIM. L. 547 (2005) (discussing the problem in Israeli jurisprudence); Timothy Macklem & John Gardner, *Provocation and Pluralism*, 64 MOD. L. REV. 815 (2001); Westen, *supra* note 8, at 139.

At the same time, some individual characteristics or beliefs of a defendant should not be incorporated into the "reasonable person." For example, a defendant who believes in white supremacy may become extremely angry at the sight of an inter-racial couple kissing in public. In applying the provocation defense, we would not want to ask "would a reasonable racist person have been provoked" by this conduct, but simply "would a reasonable person" have been provoked.[10]

The "reasonable person" formulation in provocation cases repeatedly flounders on this deep conceptual problem: to what extent should "reasonableness" be judged externally (from an objective perspective), and to what extent should "reasonableness" be judged internally—from the defendant's subjective perspective?

The Model Penal Code ("MPC") states that provocation should be assessed from the point of view of a reasonable person in the defendant's "situation," a "designedly ambiguous" term that does not make clear precisely which facts about the defendant should or should not be taken into account.[11] The commentary to the MPC takes the view that this issue cannot be "resolved satisfactorily by abstract definition."[12] Samuel Pillsbury states that the MPC's attempt to "join[] . . . objective reasonableness with individualized psychological assessment" suffers from "incoherence."[13]

Victoria Nourse suggests that the problem has no solution: "How many times are we to be asked about 'which characteristics count' for the reasonable person in provocation, negligence, and self-defense cases when we have long known that such a question cannot be answered?"[14] Paul Robinson states that scholars "cannot provide a principled theory by which to distinguish those characteristics with which the law should individualize from those with which it should not."[15]

This conceptual puzzle manifests itself in several types of recurring cases. First, the problem emerges in cases involving minority cultural or religious beliefs. Consider a defendant who becomes enraged—and kills—when the victim violates a strong cultural or religious taboo not shared by the broader political community.[16] In assessing the defendant's provocation claim, should the jury ask whether a "reasonable member of this minority cultural community" would become extremely angry in these circumstances, or should the jury ask whether an abstract "reasonable person" would become extremely angry?

---

10. Indeed, it seems inherently contradictory to speak of a "reasonable racist person." *See* Dressler, *supra* note 9, at 994 (noting that a racist person is "unreasonable").

11. MODEL PENAL CODE § 210.3 (AM. LAW INST. 1985).

12. MODEL PENAL CODE & COMMENTARIES § 210.3 cmt. 5, at 62–63 (AM. LAW INST., Official Draft and Revised Comments 1980).

13. Samuel H. Pillsbury, *Misunderstanding Provocation*, 43 U. MICH. J.L. REFORM 143, 162 (2009).

14. V.F. Nourse, *Reconceptualizing Criminal Law Defenses*, 151 U. PA. L. REV. 1691, 1733 (2003).

15. Paul H. Robinson, *Criminal Law Scholarship: Three Illusions*, 2 THEORETICAL INQUIRIES L. 287, 320 (2001).

16. *See infra* pp. 420–22 and notes 50–52.

Second, the problem emerges in one of the most common types of provocation cases—men who kill their female intimate partners.[17] Feminists have rightly criticized the provocation doctrine for, at least in some cases, mitigating the conduct of men who abuse and kill their intimate partners by re-casting the woman's efforts to separate from her abuser as acts of "provocation."[18] Yet the doctrine has not been abolished, and many commentators—including some of these critics—believe that the provocation doctrine should be available in at least some of these cases.[19] A related problem involves so-called "gay panic" or "trans panic" provocation claims, in which the defendant kills in response to a perceived romantic overture from a person who is gay, lesbian, or transgender.[20]

At first blush, these cases might seem conceptually unrelated to cases involving minority cultural or religious beliefs. At a deeper level, however, these cases likewise pose the question of whether provocation should be assessed from a perspective internal to the defendant—such as a man who is extremely possessive and jealous or a person who is repelled by same-sex intimacy—or from an objective, external perspective. The possessive man who kills his wife when she attempts to leave him does not necessarily belong to an identifiable "cultural minority," but he is asking the jury to view the reasonableness of the provocation from "his perspective," a perspective internal to his view of the world.[21] That view may include a variety of beliefs which may or may not be shared by some or all of the broader political community, such as beliefs about proper gender roles. Dan Kahan and Martha Nussbaum explain the problem as follows:

> [T]he fact of social variation creates some delicate problems when, having understood the evaluations that are internal to a given person's emotions, we

---

17. *See* Victoria Nourse, *Passion's Progress: Modern Law Reform and the Provocation Defense*, 106 YALE L.J. 1331, 1342–68 (1997).

18. *See* Donna K. Coker, *Heat of Passion and Wife Killing: Men Who Batter/Men Who Kill*, 2 S. CAL. REV. L. & WOMEN'S STUD. 71 (1992); Martha R. Mahoney, *Legal Images of Battered Women: Redefining the Issue of Separation*, 90 MICH. L. REV. 1, 78–79 (1991) (arguing that the law must recognize the phenomenon of "separation assault" in order to properly view many battering men as offenders against women's freedom rather than partial victims of female provocation); Nourse, *supra* note 17, at 1384–89.

19. *See* Dressler, *supra* note 9, at 977; Nourse, *supra* note 17, at 1337–38, 1384–89 (criticizing the doctrine from a feminist perspective yet arguing against the abolition of the provocation doctrine).

20. *See* Lee, *supra* note 8; Cynthia Lee & Peter Kwan, *The Trans Panic Defense: Masculinity, Heteronormativity, and the Murder of Transgender Women*, 66 HASTINGS L.J. 77 (2014); Robert B. Mison, Comment, *Homophobia in Manslaughter: The Homosexual Advance as Insufficient Provocation*, 80 CALIF. L. REV. 133 (1992); Victoria L. Steinberg, *A Heat of Passion Offense: Emotions and Bias in "Trans Panic" Mitigation Claims*, 25 B.C. THIRD WORLD L.J. 499, 523–24 (2005) (reviewing MARTHA C. NUSSBAUM, HIDING FROM HUMANITY: DISGUST, SHAME, AND THE LAW (2004)).

21. *See* LEE, *supra* note 9, at 110 ("Contrary to popular perception, immigrant and racial minority defendants are not the only ones who try to mitigate their criminal charges by relying on cultural norms. Every time a White heterosexual male murder defendant argues he was reasonably provoked into a heat of passion by his female partner's infidelity or by a gay man's sexual advance, he seeks to mitigate his charges by relying on cultural norms."); Ben-David, *supra* note 9, at 10 ("One of the criticisms made against recognizing cultural arguments as part of the provocation doctrine is that it improves the situation of men who killed their spouses out of jealousy.").

> then turn to the task of evaluating those evaluations for their appropriateness or reasonableness. The facts of social variation warn us that we must ask, "Whose ideas of reasonableness?" When we are asking about anger, for example, do we look for our norm of reasonableness to the Utku Eskimos or to the ancient Romans, or to some critical norm that transcends both cultures? If the answer is that we should look to our fellow-citizens and define the emotion and its norms as they do—by no means an obvious answer, since we might have good reasons to think some prevailing norms unreasonable—we still must ask, which fellow citizens, and why those?[22]

This article identifies the principle that resolves this problem: the principle of blameworthy reasons to be angry.

B. The Solution: Blameworthy Reasons to Be Angry

All of these seemingly disparate puzzles can be resolved—at a conceptual level—by a common normative principle in the heat of passion doctrine, a principle that has not yet been clearly articulated but which in a sense is embedded within the doctrine.

The principle is the following: provocation is not adequate if the reason the defendant became extremely angry is blameworthy. A belief is blameworthy if it contradicts the fundamental values of the political community. The blameworthiness principle distinguishes those features of a defendant that *cannot* form a basis for him to argue he was reasonably provoked from those features that *can* properly form the basis of a provocation claim.

A defendant who pleads provocation asks the community to mitigate his wrongful act of killing from murder to manslaughter, and to do so because another person provoked him into a rage that made it much more difficult to control his violent response. But if the defendant's reason for becoming angry is blameworthy, he has no rightful claim to mitigation—he is culpably responsible for entering the state of extreme anger that might otherwise reduce his culpability. As explained in greater detail below, this general standard solves the conceptual problem of individualizing the reasonable person in provocation cases. It provides the heretofore elusive "principled theory by which to distinguish those characteristics with which the law should individualize from those with which it should not."[23]

In cases of minority religious or cultural beliefs, it is sometimes appropriate to assess the reasonableness of the provocation from the internal perspective of that minority religious or cultural belief system. In other cases, it is not appropriate to take that internal perspective. The way to differentiate those cases is by assessing

---

22. Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 297 (1996).

23. Robinson, *supra* note 15, at 320.

whether the belief in question is blameworthy—whether it contradicts the fundamental values of the political community.

A white supremacist who is provoked by seeing an inter-racial couple kiss in public is provoked for a blameworthy reason. The belief that blacks are inferior to whites and that they should be kept separate is a belief that itself is blameworthy because it fails to "manifest appropriate respect . . . for the interests of others."[24] The white supremacist is asking the community to mitigate his wrongful act of killing from murder to manslaughter, and to do so because another person provoked him into a rage that made it much more difficult for him to control his violent response. But the white supremacist's reason for becoming extremely angry is itself blameworthy, and thus he has no rightful claim to mitigation; he is culpably responsible for entering the state of extreme anger that might otherwise reduce his culpability.

A defendant who relies on a cultural or religious belief that is not blameworthy stands in a different position. Consider the famous case of *Regina v. Gibson*, an Aboriginal defendant who killed another tribesman for violating a tribal taboo and "talking men's talk" before uninitiated boys.[25] The Aboriginal defendant is asking the community to mitigate his wrongful act of killing from murder to manslaughter, and to do so because another person provoked him into a rage that made it much more difficult for him to control his violent response. Unlike the white supremacist, it appears that the Aboriginal man's tribal taboo—against "talking men's talk" in front of uninitiated boys—is not blameworthy. It may not be a belief that the broader political community itself shares or endorses, but neither does it conflict with any of the broader community's fundamental values. Thus, the Aboriginal man may have a claim for mitigation, for he is not culpably responsible for entering the state of extreme anger that itself reduces his level of culpability.

The concept of "blameworthy reasons to be angry" also provides the correct normative standard for assessing controversial cases such as men killing their intimate female partners or cases of so-called "gay panic" or "trans panic." When defendants in these cases argue that they were provoked, and that their extreme anger was reasonable, the law should respond by determining whether (and to what extent) their anger is based on blameworthy beliefs or attitudes — attitudes that fail to manifest appropriate respect for gay or trans persons and violate the political community's commitment to gender and sex equality.

In some of these cases, a jury may determine that the defendant was partly motivated by blameworthy beliefs or attitudes, and partly by beliefs or attitudes that are not blameworthy. The infamous case of *Bedder v. Director of Public Prosecutions*,[26] discussed below in Part VI, may be one such case. These "mixed"

---

24. *But see* Westen, supra note 8, at 140.

25. *See* JOSHUA DRESSLER, CASES AND MATERIALS ON CRIMINAL LAW 283 (5th ed. 2009) (discussing *Regina v. Gibson* (unpublished opinion)); *infra* note 50.

26. [1954] 2 All ER 801, 1 WLR 1119 (UK).

cases constitute some of the most difficult provocation cases for courts and commentators to properly assess. Commentators viewing such cases from one angle have sympathy for the defendant (and wish to permit the provocation claim), whereas commentators viewing such cases from another angle condemn the defendant (and wish to deny the provocation claim).[27]

The concept of blameworthiness allows us to consider both of these perspectives and reach a principled judgment. To the extent that part of the defendant's motivation is blameworthy and part of it is not, we should exclude the blameworthy part and evaluate the sufficiency of the provocation based solely on the non-blameworthy reasons that the defendant became extremely angry.

This general principle distinguishes, at a conceptual level, those features of a defendant that are properly relevant from those that are not properly relevant. Of course, the general principle of "blameworthiness" will itself be subject to disagreement in a diverse, pluralistic society. The principle will not yield a simple or obvious answer in every case, and reasonable people may disagree about when the reason (including the belief or attribute that underlies the reason) was itself blameworthy. But "blameworthiness" does provide the correct normative standard within which these debates should occur.

Related is the question of mechanics—disagreements about what counts as "blameworthy" might be assigned to legislatures, judges, or juries (or a mix of all of the above). For example, legislatures could step in to define at least some of the reasons that are blameworthy and unacceptable, such as homophobia, racism, religious prejudice, and sexism. Alternatively, legislatures could leave the job of assessing blameworthiness entirely to juries, or judges could screen out clear cases, leaving debatable cases to juries.[28]

Both of these issues—disagreements over what is blameworthy and which legal institutions are best suited to resolving those disagreements—are substantial, and I do not claim to resolve them here.

Nevertheless, the "blameworthiness" standard resolves, at a conceptual level, the problem of which traits of the defendant should or should not be considered. It properly directs commentators, judges, and juries away from the confounding question—which characteristics of the defendant should be imported onto the "reasonable person"—and toward the correct question—is this particular defendant's reason for becoming extremely angry blameworthy?

Finally, much ink has been spilled debating whether the provocation partial defense is an excuse defense, a justification defense, or a partial justification and partial excuse.[29] For purposes of this article, it is not necessary to take a position on

---

27. *See* Nourse, *supra* note 17, at 1402–03 (discussing both views).

28. *See* Lee, *supra* note 8, at 549–57 (arguing that implicit bias is best combatted by permitting the issues to be argued and aired before the jury, rather than having the legislature or judge prohibit the jury from considering controversial issues).

29. *See* Berman & Farrell, *supra* note 2, at 1046–65 (describing the literature).

this debate. Commentators who take each of these views nevertheless seem to recognize that the "reasonable provocation" component requires a normative assessment of the defendant's emotional response,[30] and it is that normative assessment that is analyzed here. Whether that normative assessment is in the nature of an excuse,[31] a justification,[32] both excuse and justification,[33] or something else,[34] is unnecessary to resolve for present purposes.

II. HEAT OF PASSION DOCTRINE AND THE COMPONENT OF SUFFICIENT PROVOCATION

As noted above, the heat of passion doctrine contains two basic components. First, the defendant "must have killed while in a *heat of passion*" or state of extreme emotional disturbance. Second, the heat of passion or emotional disturbance must have resulted from "adequate provocation."[35]

The first component is a relatively straightforward factual inquiry, although different jurisdictions use different standards for what types of emotional states qualify for this defense. Common formulations are that the defendant must be in a "heat of passion,"[36] or a state of "extreme mental or emotional disturbance."[37] The differences among these formulations are not relevant for purposes of this article, which focuses on the second—normative—component of the partial defense.

---

30. *See infra* note 423–437.

31. *See* Joshua Dressler, *Provocation: Partial Justification or Partial Excuse?*, 51 MOD. L. REV. 467 (1988); Dressler, *supra* note 9; Peter Westen, *How Not to Argue That Reasonable Provocation Is an Excuse*, 43 U. MICH. J.L. REFORM 175 (2009).

32. *See* Berman & Farrell, *supra* note 2, at 1055 n.109 (listing articles in support of the "partial justification" view).

33. *See id.* at 1033–35.

34. *See* Kahan & Nussbaum, *supra* note 22, at 318–21 (contending that the traditional categories of excuse and justification both fail to adequately explain the provocation doctrine, as they presuppose a mechanistic conception of emotion).

35. *See* Berman & Farrell, *supra* note 2, at 1040. Courts and commentators often list more than two elements, but these alternative formulations reflect the same fundamental approach. For example, Wayne LaFave states that the defendant must establish the following elements:

> (1) There must have been a reasonable provocation.
> (2) The defendant must have been in fact provoked.
> (3) A reasonable person so provoked would not have cooled off in the interval of time between the provocation and the delivery of the fatal blow. . . .
> (4) [T]he defendant must not in fact have cooled off during that interval.

LAFAVE, *supra* note 1, at 1028; *see also* DRESSLER, *supra* note 1, at 530 (listing four elements).

This language reflects the same fundamental principles as stated in the two-part test above. Elements (2) and (4) in the LaFave formulation are both implicitly present in the first element above, that the actor must have killed while in the heat of passion. Killing while in the heat of passion implies that the defendant was in fact provoked into a heat of passion, and that he had not in fact cooled off but was still in a heat of passion. Elements (1) and (3) in the LaFave formulation are both implicitly present in the second element above, that the heat of passion was brought about by adequate provocation. The notion of "adequate provocation" can encompass both the idea that the provocation was "reasonable," as well as the notion that a reasonable person would not have cooled off but would still be in a heat of passion.

36. *See* Dressler, *supra* note 3, at 431 n.102.

37. MODEL PENAL CODE § 210.3 (AM. LAW INST. 1985); *see* Dressler, *supra* note 3, at 431 & n.105.

The normative aspect of the heat of passion doctrine comes in the second component, and it is here that courts and commentators struggle to determine which of the defendant's traits should be considered and which should not. Jurisdictions commonly provide that the provocation must be "reasonable,"[38] "sufficient,"[39] "serious,"[40] or "adequate."[41] All of these terms are meant to set a normative standard. A defendant who kills due to a provocation that is not reasonable, sufficient, serious, or adequate cannot use the heat of passion defense. For present purposes, not much seems to turn on which of these terms is used—they are not clearly defined terms with consistent (or consistently different) meanings from jurisdiction to jurisdiction.

States use a variety of formulations to try to further articulate when provocation is adequate or sufficient. Some appear to ask whether a reasonable person would be incited into using deadly force, or whether the use of deadly force is a reasonable response to the provocation.[42] Ohio, for example, requires that there be "serious provocation occasioned by the victim that is reasonably sufficient to incite the person into using deadly force."[43]

Many courts and commentators have criticized this approach, noting that even when a defendant successfully pleads heat of passion, he is still adjudged guilty of manslaughter, which is a serious felony offense.[44] Thus it seems to be a mistake to inquire whether a reasonable person would have killed or whether the killing was a reasonable response. The crime of manslaughter represents the judgment that a killing in the heat of passion is not reasonable, for it remains a serious crime.[45]

Other states recognize that the use of deadly force in a manslaughter case is never reasonable, and focus instead on whether the defendant's extreme emotional

---

38. *See* N.J. STAT. ANN. § 2C:11-4(b)(2) (West 2017) ("reasonable provocation").

39. *See* MISS. CODE ANN. § 97–3–17 (2017) ("sufficient provocation"); State v. Wharton, 672 S.E.2d 786, 788 (S.C. 2009) ("sufficient legal provocation").

40. ALASKA STAT. § 11.41.115(a) (2016) ("serious provocation").

41. ARIZ. REV. STAT. ANN. § 13-1103(A)(2) (2017) ("adequate provocation").

42. *See* State v. McGuy, 841 A.2d 1109, 1113 (R.I. 2003) (assessing whether the provocation "was 'so gross' as to cause an ordinary reasonable person to lose his self-control and to use deadly force with fatal results"); W. VA. CRIMINAL LAW RESEARCH CTR., W. VA. PUB. DEF. SERVS., WEST VIRGINIA CRIMINAL JURY INSTRUCTIONS 112 cmt. 2 (6th ed. 2003) ("The term 'provocation', as it is used to reduce murder to voluntary manslaughter, consists of certain types of acts committed against the defendant which would cause a reasonable man to kill.").

43. OHIO REV. CODE ANN. § 2903.03(A) (West 2017).

44. People v. Beltran, 301 P.3d 1120, 1130 (Cal. 2013), *as modified on denial of reh'g* (Aug. 28, 2013); Dressler, *supra* note 3, at 467–68; Westen, *supra* note 8, at 155 & n.62 (citing commentators making this point).

45. There is another, more charitable way to interpret a statute like that of Ohio. Perhaps Ohio does not mean to ask jurors to determine whether the provocation would have caused a reasonable person to kill, but rather whether the provocation would cause reasonable person to feel a strong impulse to kill. *See* Kamir, *supra* note 9, at 549 (noting that under Israeli law, the doctrine depends on "whether the reasonable person, placed in the defendant's shoes, would have been *likely* to kill as a result of the provocation") (emphasis added). This is different from asking whether a reasonable person would in fact kill. The language of the Ohio statute does not clearly differentiate between these two concepts, so it is not clear whether this reading of Ohio law is correct.

reaction was reasonable.[46] Some of these states use a "reasonable person" test, asking whether the provocation "is conduct sufficient to excite an intense passion in a reasonable person."[47]

Some states focus on the actual defendant, rather than a reasonable person, and then ask whether the defendant's emotional reaction was "reasonable." Connecticut asks whether there was a "reasonable explanation or excuse" for the defendant's extreme emotional reaction, and explains that the reasonableness of the response "is to be determined from the viewpoint of a person in the defendant's situation under the circumstances as the defendant believed them to be."[48] The California Supreme Court states, "[p]rovocation is adequate only when it would render an ordinary person of average disposition 'liable to act rashly or without due deliberation and reflection, and from this passion rather than from judgment.'"[49]

Under all of these standards, the normative question remains: what makes an extreme emotional response (however defined) "reasonable"?

III. FRAMING THE PROBLEM: CONFLICTING INTUITIONS

Courts and commentators struggle to determine when a defendant's cultural or religious beliefs should be considered when evaluating the sufficiency of provocation. A number of commentators have proposed standards purporting to draw the correct boundary between beliefs that should be considered and those that should be rejected. Part III describes four specific cases designed to elicit different intuitions about whether a defendant's cultural or religious beliefs should be considered when assessing the adequacy of the provocation. Part IV then turns to standards articulated by previous commentators, and explains why each standard fails to properly differentiate between the cases described in Part III.

To test the reader's intuitions, and to evaluate tests proposed by previous commentators, the following four cases will be considered. Each example is a case in which the defendant's extreme anger is the result of him holding a particular cultural or religious belief shared by his particular religious, ethnic, or cultural

---

46. *See* People v. Steele, 47 P.3d 225, 240 (Cal. 2002) (stating that the question is whether "the facts and circumstances were sufficient to arouse the passions of the ordinarily reasonable man") (quotation and citation omitted); State v. Flick, 425 A.2d 167, 173 (Me. 1981) (finding trial court was correct in "directing the jury to decide whether it was reasonable for [the defendant] to react with extreme anger or fear").

47. 720 ILL. COMP. STAT. ANN. 5/9-2(a), (b) (West 2017). Maryland's standard is similar. "[F]or a provocation to be 'adequate,' it must be 'calculated to inflame the passion of a *reasonable* [person] and tend to cause [that person] to act for the moment from passion rather than reason.'" Dennis v. State, 661 A.2d 175, 179 (Md. Ct. Spec. App. 1995) (alterations in original).

48. CONN. GEN. STAT. ANN. § 53a-54a (West 2017). New York law is similar: "Penal Law § 125.25(1)(a) provides that it is an affirmative defense to the crime of murder in the second degree that '[t]he defendant acted under the influence of extreme emotional disturbance for which there was a reasonable explanation or excuse, the reasonableness of which is to be determined from the viewpoint of a person in the defendant's situation under the circumstances as the defendant believed them to be.'" People v. Harris, 740 N.E.2d 227, 229 (N.Y. 2000) (alteration in original).

49. *Beltran*, 301 P.3d at 1136.

group, but not widely shared by the entire political community. Some variations on these examples are mentioned in the footnotes.

1. *Aboriginal Man*. An Australian Aboriginal defendant kills another Aboriginal man after that man "talks men's talk" in front of young boys who have not yet been initiated. Both men are members of an Aboriginal tribe that contains a strong taboo against speaking about certain tribal secrets in front of uninitiated members.[50]
2. *White Supremacist*. A defendant belongs to (and was raised in) the Aryan Nation, a community of interest gathered around the idea that the Caucasian (or Aryan) race is superior to, and should be kept separate from, other racial groups. The defendant sees an inter-racial couple kissing in public, becomes enraged, and kills an African-American man who was kissing a Caucasian woman.
3. *Devout Catholic*. A devout Catholic who works at a church observes a man desecrating the church. The man urinates on the altar and decapitates the statues of holy saints in the church. The devout Catholic becomes enraged and kills the man.[51]
4. *Conservative Pakistani Father*. A conservative Pakistani man strangles his daughter after she rejects her father's arranged marriage. The man's ethnic community has a very strong social norm that daughters always show honor and deference to their fathers, including with respect to the father's choice of an arranged marriage, and that a daughter who rejects an arranged marriage brings great shame and dishonor on her family and in particular on her father.[52]

In Case 1—the Aboriginal Man—many persons, including myself, have the intuition that the Aboriginal Man's conduct should be judged by the jury from the perspective of a reasonable Aboriginal man who shares the Aboriginal cultural

---

50. This example is taken from DRESSLER, *supra* note 25, which draws the example from the actual case of Regina v. Gibson (unpublished). *See* Robert L. Misner, *The Awkward Case of Harry Gibson*, 1986 ARIZ. ST. L.J. 691 (1986).

51. This example is based on a fictional account by Roberto Bolaño. *See* ROBERTO BOLAÑO, 2666 361–81 (Natasha Wimmer trans., 2008). A sample of Bolaño's account: "Two days later, the stranger got into the church of Santa Catalina, in Colonia Lomas del Toro, late at night when the building was closed, and he urinated and defecated on the altar, as well as decapitating almost all the statues in his path. This time, the story made the national news and a reporter from *La Voz de Sonora* dubbed the attacker the Demon Penitent." *Id.* at 367.

A similar case: an antagonist throws a pigskin shoe at a devout Muslim, knowing that the assault with take on particular force because of the Muslim religious view that pig products are forbidden (haram). *See* Macklem & Gardner, *supra* note 9, at 818.

52. This example is also taken from Joshua Dressler, who draws the example from a news article describing an actual criminal case involving similar facts in Chicago. DRESSLER, *supra* note 25, at 283–84. A similar example involves a conservative Muslim defendant who killed his wife in a rage after she made a comment suggesting she might be having an affair. R. v. Humaid, [2006] 81 O.R.3d 456 (Can. Ont. C.A.). The defense introduced evidence that traditional Muslim men live in a culture that is "male dominated, preoccupied with the concept of 'family honour,' and particularly intolerant of female infidelity." Rosemary Cairns Way, *Culture, Religion and the Ordinary Person: An Essay on* R. v. Humaid, 41 OTTAWA L. REV. 1, 7 (2009).

beliefs about "talking men's talk."[53] If not—if the jury simply asked whether a reasonable person of the dominant culture (Australian, in that case, or in our context American) would be provoked into extreme passion by a man "talking men's talk" in front of uninitiated children, the answer would be an easy and obvious "No." That approach seems unsatisfactory. It seems to somehow punish the Aboriginal defendant for holding a minority cultural belief.

Of course, judging this case from the perspective of an Aboriginal man does not mean that this particular defendant will prevail, or that the fact-finder will conclude that the provocation was adequate. Reasonable people may disagree about that ultimate issue, and the resolution will likely depend on the particular facts and testimony, including more detail than is presented above about the nature of the "talking men's talk" taboo. But in terms of framing the question, many agree that the fact-finder should take into account the Aboriginal Man's cultural beliefs when assessing the adequacy of the provocation.

In Case 2, many people, including myself, have the strong intuition that the Aryan Nation adherent's conduct should *not* be judged from the perspective of a reasonable Aryan Nation adherent who shared the Aryan Nation's cultural beliefs about white supremacy.[54] Indeed, that very formulation prompts the objection that there is no such thing as a "reasonable" Aryan Nation adherent.[55] The identifying characteristic itself—belief in white supremacy—renders the Aryan Nation adherent inherently unreasonable. Instead, the jury should determine whether a reasonable person (of the dominant culture, who does not believe in white supremacy) would have been provoked into extreme anger at the sight of inter-racial kissing in public. And the answer to that question is an easy, obvious "No." Several commentators have the same intuition for a homophobic man who kills in a rage caused by a perceived sexual advance from another man.[56]

The puzzle is why the views of the Aryan Nation adherent—who holds a cultural belief not shared by the broader community—cannot be considered by the jury, whereas the views of the Aboriginal Man—who likewise holds a cultural belief not shared by the broader community—can be considered by the jury. Why do we think it is a contradiction in terms to speak of a "reasonable Aryan Nation adherent" but not a contradiction in terms to speak of a "reasonable Aboriginal Man," even though the dominant culture does not share the relevant cultural belief of either defendant?

The remaining two examples may provoke similar responses. Case 3, the Devout Catholic, seems more like the Aboriginal Man. The broad political

---

53. Dressler, *supra* note 9, at 996. In the actual case, the judge did so instruct the jury. *See* DRESSLER, *supra* note 25, at 283 (citing Misner, *supra* note 50, at 708–10).

54. *See, e.g.*, Dressler, *supra* note 9, at 994–95.

55. *See id.* at 994.

56. *See, e.g.*, Lee, *supra* note 8, at 564; Mison, *supra* note 20; J. Kelly Strader et al., *Gay Panic, Gay Victims, and the Case for Gay Shield Laws*, 36 CARDOZO L. REV. 1473, 1501–02 (2015).

community does not necessarily share the Devout Catholic's views about the sacredness of the church and its statues. That belief is surely more common (in, say, the United States) than the Aboriginal Man's views on talking men's talk (even in Australia). But it does not seem appropriate that the result should depend on whether the Devout Catholic's conduct occurred in Rhode Island (51% Catholic) or in Mississippi (4.7% Catholic).[57] Regardless of whether the broader political community is predominantly Catholic or not, it seems clear that the jury should ask whether there was adequate provocation from the perspective of a reasonable devout Catholic, and not from the perspective of a reasonable person with no particular religious beliefs. Many have similar intuitions about a devout Muslim assaulted by a pigskin shoe.[58]

For Case 4, the Conservative Pakistani Father, I have the intuition that the father who kills his daughter for refusing to consent to an arranged marriage should not be permitted to have the jury judge his extreme anger from the perspective of a "reasonable conservative Pakistani man," but should be judged from the perspective of a reasonable person who holds appropriate values about individual autonomy and sexual and gender equality. The Court of Appeal for Ontario reasoned just this way in the case of Adi Humaid, a defendant in a Canadian case who killed his wife after she made a comment that he interpreted as suggesting she was having an affair.[59] The defense called an expert who "described Islamic culture as male dominated, preoccupied with the concept of 'family honour,' and particularly intolerant of female infidelity."[60] The Court of Appeal for Ontario recognized the troubling nature of this claim: "[T]he alleged beliefs which give the insult added gravity are premised on the notion that women are inferior to men and that violence against women is in some circumstances accepted, if not encouraged. These beliefs are antithetical to fundamental Canadian values, including gender equality."[61]

IV. ATTEMPTS TO RESOLVE THE PROBLEM: CRITIQUES

A variety of commentators have attempted to articulate standards differentiating between those cases in which a defendant's individual characteristics or beliefs should be taken into account in assessing the reasonableness of provocation and those in which the defendant's individual characteristics or beliefs should not. Part IV considers various proposals and explains how each falls short of providing a convincing general account.

---

57. *See* John Green, *State-by-State Percentage of White Evangelicals, Catholics, and Black Protestants*, BELIEFNET, https://www.beliefnet.com/Faiths/2004/11/State-By-State-Percentage-Of-White-Evangelicals-Catholics-And-Black-Protestants.aspx (last visited Feb. 15, 2018).

58. *See* Macklem & Gardner, *supra* note 9, at 818.

59. Way, *supra* note 52, at 7–8.

60. *Id.* at 7.

61. R. v. Humaid, [2006] 81 O.R. 3d 456, 480 para. 93 (Can. Ont. C.A.); *see also* Way, *supra* note 52, at 11.

A. Gravity of the Provocation and Self-Control

Andrew Ashworth addresses how to "individualize" the abstract "reasonable man" in provocation cases.[62] He argues that under the provocation doctrine, "[t]he law's paramount concern is to ascertain whether the accused showed a reasonable amount of self-restraint."[63]

Ashworth notes that the "reasonable man" test conflates two issues.[64] First, the term "reasonable" seems to suggest a judgment of "reasonable self-control."[65] Second, the phrase "reasonable man" "begs questions about the age, sex, marital status, race, colour, religion and other personal characteristics of a hypothetical individual."[66]

Ashworth notes that many individual characteristics of the defendant must be taken into account to make sense of the "provocation."[67] The sight of two persons engaging in sexual intercourse is not "grave provocation," but it might be for someone who is married to one of the participants.[68] Likewise, "to say that throwing a pigskin shoe is a grave provocation would be incorrect as a general proposition," but it might be a grave provocation given that both persons involved were Muslims, which substantially magnifies the degree of provocation.[69]

Thus, we must assess "the 'gravity' of provocation . . . in relation to persons in a particular situation or group."[70] Accordingly, Ashworth argues, "it is essential and inevitable that the accused's personal characteristics should be considered by the court."[71]

At the same time, Ashworth argues, the court must also disregard certain traits of the defendant. "If it is accepted that a primary purpose of the 'reasonable man' test is to ascertain whether the accused showed reasonable self-control in the face of the provocation given, then it follows that individual deficiencies of temperament and mentality must be left out of [the] account."[72]

Ashworth thus recognizes that courts must make a normative judgment about the appropriateness of the defendant's response. To make that normative judgment, Ashworth proposes the following rule: "The proper distinction, it is submitted, is that individual peculiarities which bear on the gravity of the provocation should be taken into account, whereas individual peculiarities bearing on the accused's level

62. A.J. Ashworth, *The Doctrine of Provocation*, 35 CAMBRIDGE L.J. 292, 299 (1976).
63. *Id.*
64. *Id.*
65. *Id.*
66. *Id.*
67. *Id.* at 300.
68. *Id.*
69. *Id.*
70. *Id.*
71. *Id.*
72. *Id.*

of self-control should not."[73]

Ashworth's rejection of "individual peculiarities bearing on the accused's level of self-control" seems sound. The concept of "reasonable" or "adequate" provocation suggests a judgment about whether the defendant's response was in some way proper. If the reason for the extreme anger is itself a feature of the defendant that demonstrates improper sensitivity for the interests of others, then the provocation was not "reasonable." Thus, a defendant who lacks self-control can be faulted for exactly that feature of his character.

The first part of Ashworth's test, however, is too broad. Ashworth suggests that "individual peculiarities which bear on the gravity of the provocation should be taken into account." Consider the examples in Part III under this standard. Both the White Supremacist and the homophobic man have "individual peculiarities"—a strong belief in white supremacy, and homophobia, respectively—which "bear on the gravity of the provocation." The provocation of the White Supremacist does not necessarily stem from his level of self-control. He might be a relatively cool-tempered person as a general matter. If so, it is not his short temper, but his deep moral commitment to white supremacy, that explains his response.

Accordingly, there are some cases (such as the White Supremacist, the Conservative Pakistani Father, and the homophobic man) in which the defendant's "individual peculiarities which bear on the gravity of the provocation"[74] do not seem properly relevant. Ashworth's test fails to allow us to reject the claims of those defendants.

B. Insults That "Really Are Insulting"

Timothy Macklem and John Gardner ask, "What is the best way to reflect human diversity in the structure of the provocation defence . . . ?"[75] They helpfully break the problem down into three distinct inquiries by proposing the following test: "first, was there an action capable of constituting a provocation? Second, how provocative was it? And third, how much self-control should have been exhibited in the face of it?"[76]

To answer the first question, Macklem and Gardner note that "courts have clearly been right to warn juries in provocation cases that different words and deeds have different significance for different defendants."[77] As Ashworth also recognized, this is because the defendant "might inhabit a different social milieu from the jury . . . , and so might participate in a different menu of possible slights and put-downs."[78] For example, "[t]he fact that the defendant is a Muslim, say, can

73. *Id.*
74. *Id.*
75. Macklem & Gardner, *supra* note 9, at 815.
76. *Id.*
77. *Id.* at 818.
78. *Id.*; *see also* Ashworth, *supra* note 62, at 300.

be pertinent to the question of whether there was a provocation only if there is a Muslim social milieu in which there is a distinctive menu of possible insults."[79]

After determining whether there was some action "capable of constituting a provocation," the second question is "whether she really *should* have been provoked by it as she was."[80] This raises two subsidiary issues. The first is "whether the things that are intelligible as insults in a particular cultural milieu really are insulting."[81] The second issue is "how serious (or grave) an insult a certain insult is."[82]

Ultimately, however, their analysis fails to adequately allow courts and juries to distinguish cases appropriately. The problem with Macklem and Gardner's scheme is how we are to assess whether a particular insult "really is insulting." At this point in the analysis, it is clear that Macklem and Gardner are imposing a normative standard. For example, the "question of whether being called 'queer'" comes down to the "question of whether it really is an insult at all."[83] This depends, they argue, on "whether [the defendant] should have been angered by it at all."[84] Later they explain: "[i]t is not a matter of *why* one lost one's temper, but a matter of why one *should* have lost it."[85] "Should" is the operative word here, demonstrating that whether an insult "really is an insult at all" depends on applying some normative assessment of what properly counts as an insult. Macklem and Gardner thus argue:

> These remarks expose the limited way in which a defendant's personal idiosyncrasies can bear on the gravity of the provocation. It is not that he has his own judgment of what is insulting, . . . a judgment that needs somehow to be respected or accommodated by the law. His judgment in these matters is no more authoritative than that of his or any other community: he and all his peers might wrongly think that homosexuality is obnoxious and so might become quite unjustifiably enraged.[86]

Accordingly, "[t]he mere fact that the parties or their peers or indeed all the inhabitants of their social milieu *think* the words or deeds insulting, . . . does not even begin to make them so."[87] Instead, "[a]n insult is only as insulting as it really is, when addressed to this defendant in these circumstances. No amount of thinking it more insulting, on the part of the defendant or his peers or society at large, can

79. Macklem & Gardner, *supra* note 9, at 818.
80. *Id.* at 820.
81. *Id.* at 821.
82. *Id.*
83. *Id.*
84. *Id.*
85. *Id.* at 824.
86. *Id.* at 822–23.
87. *Id.* at 823.

make it more insulting."[88]

The notion that a particular insult "really is insulting" (or "really is not insulting") suggests that the insult must be evaluated from some objective standpoint—at least, from a standpoint removed from the internal views of the culture or personal beliefs of the defendant. But Macklem and Gardner do not explain what that objective standpoint should be, or how it applies.

Consider a Muslim defendant. In one case, the defendant becomes enraged and kills his wife after hearing a suggestion that she might be having an affair with another man.[89] Is this "really" a grave insult or not, given testimony claiming that traditional Muslim men are extremely sensitive to female infidelity?

Compare that Muslim defendant to a Muslim defendant who becomes enraged and kills after being assaulted by a person who throws a shoe made of pigskin, knowing that the Muslim man will experience additional insult and anger due to the violation of his religious prohibition against pigskin. Is that "really" a grave insult or not, given testimony about the Muslim religious prohibition?

Many would like to distinguish the first case from the second. Macklem and Gardner do not provide a standard for resolving this normative judgment, but only illustrate that there is a judgment to be made.

C. Reasons Shared by the Political Community

Victoria Nourse and Peter Westen have separately addressed this problem and each proposed a test to resolve it. The two tests are distinct, but they share a common approach: both ask whether the defendant's reason for becoming extremely angry is a reason shared or endorsed by the broader political community.[90]

Nourse relays the following story: "After days of deliberation in a case in which a defendant killed a man who had parked in his parking place, one jury summed up its confusion about the [provocation] defense by sending a note to the judge, asking, Whose norms apply, his or ours?"[91] Nourse and Westen each (in their own way) resoundingly reply: "Ours!"

This common approach avoids the problem created by Ashworth's test, in that it denies the provocation defense to individuals—like the White Supremacist—who become enraged due to certain objectionable personal beliefs.

But both Nourse's test and Westen's test fail to allow a jury to distinguish between a case like White Supremacist and a case like the Aboriginal Man. Under their respective tests, the law would deny the Aboriginal Man access to the provocation partial defense, because the broader political community likely does

---

88. *Id.*

89. *See* R. v. Humaid, [2006] 81 O.R. 3d 456 (Can. Ont. C.A.).

90. *See also* LEE, *supra* note 9, at 238 ("A more difficult inquiry, but one which better addresses the question of reasonableness as something more than whatever happens to be popular, is whether the judge or jury ought to *give society's stamp of approval* to the defendant's actions.") (emphasis added).

91. Nourse, *supra* note 17, at 1372.

not share or endorse the Aboriginal Man's reason for becoming extremely angry—his response to a violation of the "talking men's talk" taboo. The same seems true for the Devout Catholic, or the Muslim Man assaulted by a pigskin shoe.

Nourse argues for the following standard: "[t]o merit the reduction of verdict typically associated with manslaughter, the defendant's claim to our compassion must put him in a position of normative equality vis-à-vis his victim."[92] This means, Nourse explains, that the defense should be available "when the defendant appeals to the very emotions to which the state appeals to rationalize its own use of violence."[93] In other words, courts should look to the criminal prohibitions in the jurisdiction—which are the circumstances in which the state claims proper authority to apply criminal punishment to its citizens—to assess permissible reasons for becoming provoked.[94]

Nourse uses a variety of examples to illustrate how this standard works. Two of her examples, briefly stated, are as follows:

1. *Rapist Killer*. A woman is raped by Rapist. After the rape is complete, and when Rapist is departing, the woman kills Rapist in a fit of rage.[95]
2. *Departing Wife Killer*. A man kills his wife in a fit of rage after she announces that she is leaving him and will seek a divorce.[96]

In the case of the Rapist Killer, Nourse argues that "we feel 'with' the killer because she is expressing outrage in ways that communicate an emotional judgment (about the wrongfulness of rape) that is uncontroversially shared, indeed, that the law itself recognizes."[97] The defendant's "sense of emotional wrongfulness" is reflected in "the law's own sense of appropriate retribution."[98] "In this sense, the defendant is us."[99] Thus, Nourse is claiming that provocation may be "sufficient" or "reasonable" in cases in which the community (here, as reflected in the law) shares the defendant's judgment about the rightness of her anger.

---

92. *Id.* at 1396. Earlier Nourse states the test as follows: "The passion defense should be retained as a partial excuse but only in the limited set of cases in which the defendant and the victim stand on an equal emotional and normative plane." *Id.* at 1337.

93. *Id.* at 1338.

94. *Id.* at 1392–93.

95. *Id.* at 1390–91. A rape victim who kills before or during rape will likely have a valid self-defense claim, as the victim of a rape almost surely faces (or reasonably believes she faces) an imminent threat of death or serious bodily harm. But if the rapist is departing, that imminent threat may no longer exist, and self-defense may not apply. Nourse's "Rapist Killer" hypothetical appears to be designed to suggest that the victim is not acting in self-defense, but may still successfully plead the heat of passion partial defense.

96. *Id.*

97. *Id.* at 1392.

98. *Id.*

99. *Id.*

In the case of the Departing Wife Killer, the defendant "asks us to share in the idea that leaving merits outrage."[100] But that is not a claim the community—as reflected in the law—shares. On the contrary, "the law tells us quite the opposite: that departure, unlike rape and battery and robbery, merits protection rather than punishment."[101]

Nourse argues that provocation is sufficient when a defendant "respond[s] with a rage shared by the law."[102] In these cases, the defendant is not simply making "a claim for sympathy," but rather making "a claim of authority and a demand for our concurrence."[103] The defendant "asks that we share his judgments of emotional blame."[104]

Like Nourse, Peter Westen argues that the fact-finder must not adopt the defendant's internal values about when it is right to become extremely angry, but must make its own normative judgment. Also like Nourse, Westen asks whether the community shares or endorses the defendant's rage. Beyond those points of agreement, Westen's proposal differs from Nourse's.

Westen argues that the reasonableness of provocation should be assessed by a normative judgment involved in much of criminal law—the defendant's blameworthiness.[105]

Westen's first insight is that courts and commentators are starting with the wrong question: "[t]he question is not, 'To what extent should an idealized 'reasonable person' be 'individualized' with a defendant's individual traits (or 'contextualized,' 'subjectivized' or 'relativized,' as some commentators put it). . . ."[106] This question mistakenly suggests that the answer lies in figuring out which of an individual's traits are relevant and which are not relevant, as compared with an abstract "reasonable person." "[T]his effort to solve a problem of culpability by focusing on certain physical, psychological and emotional traits shifts attention away from the essential nature of blame itself."[107]

Westen argues the courts should not attempt to "individualize" the abstract reasonable person with some of the defendant's traits and features, but rather should begin with the actual defendant and then apply the appropriate moral filter to evaluate that particular defendant's reaction.[108]

---

100. *Id.*
101. *Id.*
102. *Id.* at 1393.
103. *Id.*
104. *Id.*
105. Westen, *supra* note 8, at 151.
106. *Id.* at 140.
107. *Id.* at 148.
108. *Id.* at 140. Nourse, in effect, seems to be making the same basic move—attempting to identify the appropriate moral filter to apply to the defendant's extreme anger, rather than focus on which characteristics of the defendant should or should not be imported onto the reasonable person.

The reasonableness of a defendant's extreme anger—that is to say, the reasonableness of the provocation—depends on the "actor's blameworthiness."[109] Declaring an actor blameworthy, Westen explains, "is to adjudge that, rather than being motivated in his conduct by proper regard for interests that the law seeks to safeguard, the person placed insufficient value on those interests."[110]

Accordingly, Westen proposes that blameworthiness be assessed as follows: "What would a person, who otherwise possessed every trait of the actor but fully respected the interests that the statute at hand seeks to protect, have thought and/or felt on the occasion at issue?"[111]

Translating this into the specific context of the heat of passion doctrine, Westen argues that the key question is, "Given everything that makes a defendant who he empirically is—that is, every physical trait, historical experience, and emotional disposition that constitutes him—does his inadvertence or extreme emotional agitation nevertheless manifest appropriate respect on his part for the interests of others?"[112]

Westen also proposes jury instructions designed to enable a jury to take into account all relevant parts of the defendant's background, experience, and personal traits, while excluding certain parts of the defendant's beliefs. Namely, Westen tells the jury to disregard the defendant's personal views about what it is right to be extremely angry about and substitute the judgment of the community as to that question.

Westen proposes the following instruction:

> *The Defendant's Traits*. To [determine whether the defendant's extreme anger was reasonable], you shall start by taking the defendant as he is—with one important exception that I'll discuss in a moment. That is, you shall consider everything you know about the defendant—including every physical characteristic, every experience, every personal relationship to others, and every emotional and temperamental disposition he may have.
>
> *The Defendant's Values*. Nevertheless, there is one thing about the defendant you shall *not* take into account. You shall not consider his own personal views, or the particular views of his ethnic community, about what it is right to be extremely angry about. Rather, you shall judge the reasonableness of his anger by how angry you, as representatives of the people of the state, conclude an individual who is otherwise like him is *right* to feel—or, at least, *not wrong to feel*—in response to what actually upset him.[113]

---

109. *Id.* at 151.
110. *Compare id.*, *with* Nourse, *supra* note 17, at 1393–94 (similar idea).
111. Westen, *supra* note 8, at 151.
112. *Id.* at 140.
113. *Id.* at 158.

The descriptions above show that Nourse's test and Westen's test are conceptually distinct.[114] Both tests, however, share a fundamental similarity. Both allow the provocation partial defense only in cases in which the broader political community affirmatively shares or endorses the defendant's anger in response to the provocation. Both Westen's test and Nourse's test fail, however, to adequately differentiate the test cases stated in Part III. That is to say, neither test allows us to differentiate between cases in which a defendant's personal or cultural beliefs do seem properly relevant, and cases in which they do not.

Return to the Aboriginal Man. Under Westen's proposed jury instruction, the jury cannot consider the fact that the Aboriginal culture (of which the defendant and victim were both members) contains a strong taboo against "talking men's talk." Westen's instruction explicitly tells the jury it may not consider the fact that the Aboriginal culture believes it is right (or at least not wrong) to become extremely angry in response to a violation of its taboo against "talking men's talk." That is to say, the jury is prohibited from considering the belief in the taboo—because belief in the taboo and belief that violating the taboo is a good cause for extreme anger are one in the same. Under this instruction, then, the Aboriginal defendant will fail in his "provocation" claim, unless the jury is composed of members of the same Aboriginal tribe and the jurors share the same belief in the "talking men's talk" taboo.

Now, consider the White Supremacist. Westen's jury instruction, like in the case of the Aboriginal Man, tells the jury that it cannot consider the defendant's beliefs about what it is right to be angry about. For the White Supremacist, that means that the jury cannot consider the fact that the defendant (and perhaps his sub-culture) strongly believes that inter-racial romance is wrong, and that the defendant believes that public displays of inter-racial affection are a good reason to become extremely angry.

For the White Supremacist, that result aligns with our intuitions. We do not want to judge this defendant from the point of view of someone who believes inter-racial romance is wrongful. We prefer to judge him by the norms of the community—which endorses norms of racial equality and non-discrimination.

But for the Aboriginal Man, the result seems perverse. It seems appropriate to judge the defendant according to the norms of the community to which he and the victim both belong, even though those norms may not be broadly shared by the larger political community (or the jury representing that community).

Westen's instruction likewise fails to differentiate between the Conservative Pakistani Father and the Devout Catholic. Westen's formulation rules out the heat

---

114. Nourse's test asks if the cause for the defendant's anger is one already recognized by the state as a crime; Westen's test does not require this, but instead asks more generally whether the community shares or endorses the defendant's anger.

of passion defense in any case in which the defendant's reason for becoming angry is one that is not broadly shared by the wider political community.

The flaw in Westen's proposed instruction is that it limits "reasonable provocations" to provocations for which the broader community all agree form a good reason to be extremely angry. But there are some cases, such as the Aboriginal Man or the Devout Catholic, in which the broader community may not *share* the defendant's outrage at the provocation, and yet many feel the defense should be available.[115]

Nourse's test suffers from the same flaw. Nourse asks whether the defendant "respond[ed] with a rage shared by the law."[116] If so, the defendant is not simply making "a claim for sympathy," but rather making "a claim of authority and a demand for our concurrence."[117] The defendant "asks that we share his judgments of emotional blame."[118]

The Aboriginal Man loses under this test. The law of the jurisdiction does not share—or reflect—his rage at a fellow Aboriginal tribesman who "talks men's talk" in front of an uninitiated boy. Not only does Nourse require (like Westen) that the community "share" the defendant's judgment "of emotional blame," but she further requires that this judgment be reflected in the law. For a defendant like the Aboriginal Man, whose anger derives from a deeply held cultural conviction (even if it is a conviction shared by the victim), the provocation doctrine is not available.

The Devout Catholic is more complicated to evaluate under Nourse's test, although ultimately this example illustrates an additional flaw in Nourse's standard. Unlike the Aboriginal Man—who can point only to a violation of deeply held cultural norms—the Devout Catholic can also point to a violation of the criminal law. The Devout Catholic responds to a man who urinates on the altar of a Catholic church and decapitates holy statues. Those acts—certainly the decapitation of the holy statues—are property crimes. If the statues are very valuable, they may be somewhat serious property crimes. Thus, the Devout Catholic might argue that his anger is shared by the political community—a judgment reflected by the criminal law's prohibition against damaging the property of another.

Yet, ultimately, the question remains whether the fact-finder should view the desecrations of the sacred Catholic statues from the perspective of a Devout Catholic, or from the perspective of a neutral "reasonable person." If the Devout Catholic is fortunate enough to draw a jury composed predominantly of devout Catholics, perhaps he will prevail even under a "generic" "reasonable person" standard—those jurors may well embody the "reasonable person" with their own views of sacredness. But this is just a fortuity. The Devout Catholic may

115. *See infra* pp. 420–22 and notes 53–58.
116. Nourse, *supra* note 17, at 1393.
117. *Id.*
118. *Id.*

likewise find himself before a jury (or judge) that does not place any particular value on Catholic statuary.

Moreover, the deep offense that the Devout Catholic seeks to have the jury recognize does not come primarily from the fact that the transgressor damaged property, or that the criminal law punishes deliberate acts of property damage. Would a reasonable person be provoked into extreme anger—the sort of anger that makes it much more difficult to restrain an impulse to kill—by seeing a man damage property in a public park? Probably not. The Devout Catholic's strongest argument does not come from the fact that property was damaged—although that is certainly relevant—but from the fact that the defendant deliberately chose to desecrate (not merely damage) property held to be deeply sacred by the Devout Catholic and his church.

Nourse would permit the jury to consider the criminal act of property damage as a potentially sufficient provocation, but it does not appear under her test that the jury should also consider the defendant's deliberate choice to lewdly desecrate a sacred and holy object, which is precisely the reason this instance of property damage is so enraging. The same is true for the Muslim Man assaulted by a pigskin shoe.

Both Nourse and Westen require that the broader political community affirmatively share or endorse the defendant's reason for becoming extremely angry. Accordingly, both tests fail to accommodate defendants like the Aboriginal Man or the Devout Catholic.

D. The "Oxymoron Principle" and the "Anti-Violence" Norm

Joshua Dressler has also directly addressed the problem of how much to "subjectivize" or "individualize" the reasonable person in provocation cases. He concedes that he is "unsure of precisely where . . . to draw the lines in specific cases,"[119] but he offers two rules to distinguish these cases.

The first he labels the "oxymoron principle."[120] He explains that a jury should not incorporate "an attribute or belief of the provoked actor that would force us to describe the 'reasonable person' in near oxymoronic terms. One cannot be a 'reasonable paranoid,' for example."[121] Likewise there is no such thing as a "reasonable white supremacist," so the jury should not be instructed to view the sufficiency of the provocation from the perspective of a "reasonable white supremacist."[122]

Second, if a defendant argues provocation based on a cultural belief system not shared by the broader culture, "we must decide if allowing him to explain his

119. Dressler, *supra* note 9, at 994.
120. *Id.*
121. *Id.*
122. *Id.*

reason for his anger on the basis of a belief system foreign to the jury would violate the anti-violence norms of the criminal law."[123] Dressler uses the example of the Aboriginal Man, and concludes that his provocation claim should go to the jury.[124] The cultural belief motivating his extreme anger does not violate the anti-violence norms of the criminal law—the cultural taboo against "talking men's talk" does not seem to "suggest[] that Aboriginals value human life less than non-Aboriginals."[125]

Dressler contrasts the Aboriginal Man with individuals who have moral norms that do violate the criminal law's anti-violence norm: for example, an assassin who believes killing others is not wrong.[126]

The "oxymoron principle" functions as follows: by juxtaposing the particular belief or attribute in question (say, homophobia) with the concept of the "reasonable person," Dressler prompts us to evaluate whether the particular belief or attribute is itself "reasonable." Asking whether a "reasonable white supremacist" is an oxymoron is a way of prompting us to evaluate the "reasonableness" of believing in white supremacy.

This test is thus similar to the "blameworthy" test I articulate below, although the "blameworthy" test makes explicit what is merely implied by Dressler's "oxymoron" test. Dressler's "oxymoron" principle allows us to distinguish between the Aboriginal Man and the White Supremacist. There is nothing oxymoronic about referring to a "reasonable Aboriginal man," but the concept of a "reasonable white supremacist" is an oxymoron.

The "oxymoron" principle becomes more difficult to apply in the case of the Conservative Pakistani Father, or the Adi Humaid case. Is it an oxymoron to refer to a "reasonable Pakistani man"? Surely not. Is it an oxymoron to refer to a "reasonable conservative Pakistani man"? Again, the answer seems to be no. Perhaps the "oxymoron" test should be applied at a greater level of specificity: is it an oxymoron to refer to a "reasonable conservative Pakistani man who believes that daughters should submit to their father's choice of husbands"? Or consider *R. v. Humaid*. Do we ask whether a "reasonable Muslim man" is an oxymoron? A "reasonable Muslim man whose culture is extremely intolerant of female infidelity"?

These examples begin to show the shortcomings of the "oxymoron" test. It can be difficult to ensure that the test is being applied at the right level of specificity, or targeted to the appropriate attribute or belief. More fundamentally, the test becomes awkward in some cases because it uses the concept of "oxymoron" to approximate what is really going on: namely, evaluating whether the particular belief or attribute in question is blameworthy.

Dressler's "oxymoron principle" functions by adopting a particular meaning of the concept of "reasonableness," but he is not explicit about what "reasonableness"

123. *Id.* at 996.
124. *Id.*
125. *Id.*
126. *Id.* at 991–92.

means. One view of what counts as "reasonable"—a view expressed by Nourse and Westen—is whether the broader political community shares or endorses the defendant's emotional response. If that is what Dressler means by "reasonable," then his view is similar to Nourse and Westen (and is flawed for the same reasons).

But Dressler does not appear to mean that, as he concludes that a "reasonable Aboriginal man" is not an oxymoron. Thus, the Aboriginal Man—and, most importantly, his cultural belief regarding the taboo against "talking men's talk"—can be viewed as "reasonable."

Accordingly, Dressler appears to be using the term "reasonable" in a particular way. While he is not explicit, his examples are consistent with the meaning of "reasonable" articulated here—that "reasonableness" relates to "blameworthiness." That is to say, a defendant's extreme emotional reaction is not "reasonable" if that reaction is blameworthy—if it conflicts with the fundamental values of the political community.

Dressler's second rule—the "anti-violence" principle—seems to simply be one manifestation of the "oxymoron principle," and thus is an unnecessary addition. Dressler invokes the "anti-violence" rule as a way to reject a provocation claim by an assassin or terrorist who simply does not value human life appropriately.[127] But the "oxymoron principle" is already sufficient to reject those claims. A defendant who asks to be judged from the perspective of a "reasonable assassin" is pressing an oxymoron, for an assassin—someone who kills without adequate justification—is unreasonable.

Accordingly, Dressler's "anti-violence" norm seems to be a subset of his "oxymoron" principle. That "oxymoron" principle points in the right direction, yet it merely implies what should be made explicit—the "blameworthiness" principle explained below.

E. Reasonableness and "Switching"

Cynthia Lee agrees with the commentators above that jurors should ask not only whether the defendant in fact became extremely angry, but "should also ask whether the defendant's response ought to be recognized as reasonable."[128]

Lee recognizes the difficulty posed by defendants with specific cultural or religious beliefs,[129] as well as the problem of how to assess reasonableness given different perspectives based on gender, race, and sexual orientation.[130] She agrees that the law must make a normative assessment of reasonableness, not merely assess descriptively whether "most people" would become angry in the situation.[131]

---

127. *Id.* at 996–97.
128. LEE, *supra* note 9, at 246.
129. *Id.* at 209–12
130. *Id.* at 212–25.
131. *Id.* at 226–60.

The remaining difficult question, of course, is how to articulate a standard or guideline for the normative judgment jurors are asked to make—in Lee's terms, how "[t]o give descriptive content to normative reasonableness."[132] Lee invokes the concept of "switching" to enable jurors to appropriately conduct this normative inquiry.[133]

"Switching" (in this context) is the concept of asking jurors to "switch" perspectives based on attributes such as gender, race, or sexual orientation.[134] Take a stereotypical provocation case in which "the defendant is a man who claims he was provoked into a heat of passion when he caught his female partner in an act of infidelity."[135] When assessing the "reasonableness" of this provocation, jurors might switch genders of the parties involved, asking "what the average woman would do if she came home and found her male partner in bed with another woman."[136]

Lee contends that "[m]ost women in this situation would be upset, but few would become so violent as to kill."[137] Thus switching gender perspectives might help "jurors understand the unreasonableness of the male defendant's use of deadly force."[138]

Switching also could be used in cases involving race or sexual orientation. The jurors might be instructed to consider a similar set of facts with the race or sexual orientation of the parties changed.[139] This act of switching, Lee argues, can help reveal hidden or unspoken bias and prejudice on the part of the jurors.[140] Lee draws on social science evidence showing that when hidden biases become more visible and salient, jurors actually are better able to recognize and overcome their implicit biases.[141]

> Accordingly, Lee proposes
> that judges instruct jurors that while it is completely normal to be influenced by dominant social norms, including masculinity norms,

---

132. *Id.* at 252.
133. *Id.* at 217–25, 252–53.
134. *Id.* at 217–25.
135. *Id.* at 218.
136. *Id.*
137. *Id.*
138. *Id.*
139. *Id.* at 221–25, 252–53.
140. *Id.* at 217–25, 258–59 ("Role reversal through gender-, race-, and sexual orientation-switching is a useful way of exposing bias and privilege. Switching helps expose the fact that heterosexual male violence is often presumed reasonable when similar violence by heterosexual females, gay men, and men of color would not be seen as reasonable."); Lee, *supra* note 8, at 564–66.
141. Lee, *supra* note 8, at 536 ("[R]esearch suggests that making race salient (i.e., calling attention to race) can help individuals to overcome what would otherwise be automatic stereotype-congruent responses. I argue that if making race salient makes it easier for individuals to battle racial stereotypes, then making sexual orientation salient may similarly make it easier for individuals to battle sexual orientation bias."); LEE, *supra* note 9, at 259 ("Switching is . . . likely to spark debate and discussion in the jury room about whether violence in response to female infidelity, gay panic, or racialized fear ought to be considered reasonable.").

> heterosexuality norms, and race norms, they should try not to let such norms bias their decision making. If they are uncertain as to whether such norms have influenced them, jurors may engage in gender-, race-, and/or sexual orientation-switching.[142]

While this technique is promising, Lee recognizes the potential pitfalls of using "switching" as a general technique. She states, for example, that gender-switching is not "appropriate in all cases."[143] For a "battered woman who kills her abuser during a confrontation," gender-switching would ask the jury to consider whether the man—who may be larger and physically stronger—would have responded similarly to abuse from a woman.[144]

Likewise Lee argues that race-switching would be inappropriate in a case in which the defendant, a black man, killed in response to a grave racial insult.[145] In that context, Lee argues that the jury should be educated on the history of discrimination—and the role of the "N-word"—in denigrating black persons in America.[146] It makes no sense to consider this insult from the perspective of a white person, as this insult does not denigrate white persons.[147] "Even if jurors were to do more than simply switch the races of the parties, such as replacing the N word with the word 'honky,' race-switching might still be inappropriate" because whites do not have the same history of racism, and racial slurs, as blacks.[148]

Lee's comments thus show that while "switching" can provide a useful tool for unearthing and hopefully defusing implicit bias, the practice does not actually identify the fundamental normative judgment to be made. To determine when switching is appropriate, the criminal justice system needs some standard against which to evaluate the appropriateness, or inappropriateness, of the defendant's emotional response.

V. REASONABLE PROVOCATION AND BLAMEWORTHY REASONS TO BE ANGRY

The proposals set forth above either fail to adequately distinguish our model cases, or fail to provide convincing conceptual accounts. There is general agreement that assessing the "reasonableness" or "sufficiency" of provocation involves a normative assessment of the defendant's emotional reaction, but no agreement on the exact nature of that normative assessment.[149]

---

142. LEE, *supra* note 9, at 252–53.
143. *Id.* at 219.
144. *Id.* at 219–20.
145. *Id.* at 224–25.
146. *Id.*
147. *Id.* at 225.
148. *Id.*
149. *See supra* note 423–437.

A. Normative Evaluation of Emotional Response: Blameworthy Anger

Dan Kahan and Martha Nussbaum articulate and defend an "evaluative conception" of emotion.[150] The blameworthiness principle described in this article is just that—a normative evaluation of a defendant's emotional response.

Kahan and Nussbaum's evaluative conception holds that emotions themselves are not merely mechanistic, automated responses to stimuli.[151] Instead, "emotions contain not only thought, but thought of a particular sort, namely *appraisal* or *evaluation*."[152] Because emotions themselves express appraisal or evaluations, emotions can in turn be properly subject to normative assessment or evaluation.[153] Some emotional responses are normatively appropriate or laudable; others are inappropriate or blameworthy.

In the context of the provocation doctrine, Kahan and Nussbaum explain how an evaluative conception of emotion helps illustrate the proper functioning of the doctrine:

> There must be provocation because it is only in response to significant slights that anger or rage is morally appropriate. And mitigation is warranted only when the angry person attacks her provoker because it is *that* person's conduct that forms the proper object of the evaluation embodied in anger. These features of the doctrine prevent mitigation in circumstances in which the actor's passion, however intense, reflects inappropriate valuations.[154]

An appropriate test for provocation thus will involve a normative evaluation of the defendant's extreme anger. Moreover, identifying the correct type of normative evaluation helps solve the dilemma of cultural claims.

The correct standard is that provocation is not "adequate" or "sufficient" if the defendant's reason for becoming extremely angry is itself blameworthy. This standard does not ask if the broader community endorses or shares the defendant's reason for becoming extremely angry. In cases like the Aboriginal Man, the Devout Catholic, or the Muslim Man assaulted by the pigskin shoe, the broader political community may not share or affirmatively endorse those reasons or values. Instead, this standard asks whether the broader political community condemns or affirmatively rejects the defendant's reason for becoming extremely angry—that is, whether the reason for becoming extremely angry is blameworthy.

The heat of passion doctrine recognizes that persons in a state of passion—such as extreme anger—have substantially less control over their actions, and thus they are less blameworthy (or culpable) for acting rashly than a person who does so

---

150. Kahan & Nussbaum, *supra* note 22, at 285–301.

151. *Id.* at 286.

152. *Id.*

153. *Id.* at 287–88.

154. *Id.* at 306–07; *see also* Westen, *supra* note 8, at 140 (asking whether the defendant's emotional response fails to "manifest appropriate respect . . . for the interests of others").

while in full control of their emotions. At the same time, some persons in a state of extreme anger are themselves to blame for being in that state. Those persons are excluded from claims of reduced culpability precisely because they are to blame for placing themselves in a state of reduced control, and thus cannot ask our sympathy when they act rashly in that condition.

B. Blameworthy Anger—Distinguishing Disputed Cases

An example of a proper jury instruction for the "reasonable provocation" component of the provocation doctrine is the following:

> *Reasonableness of Anger.*
>
> You must determine whether it was reasonable for the defendant to become extremely angry. To determine this, you must begin by looking at things from the defendant's perspective, taking into account everything you know about the defendant—including every physical characteristic, experience, character trait, and cultural, religious, or personal belief.
>
> A defendant's extreme anger is not "reasonable" if it is based on some blameworthy belief, attitude, or trait. The reason is blameworthy if it conflicts with the fundamental values of our nation. You must determine whether the reason that the defendant became extremely angry is a blameworthy reason.
>
> If you determine that the defendant's extreme anger does not come from some blameworthy belief, attitude, or trait, then you shall determine whether it was reasonable for a person in the defendant's position to become extremely angry.
>
> If you determine that the defendant's extreme anger came partly from a blameworthy reason (or reasons), and partly from a reason (or reasons) that is not blameworthy, then you shall ignore the blameworthy reason(s) and focus only on the reasons that are not blameworthy. Considering only the non-blameworthy reasons, you shall determine whether, in light of those reasons, it was reasonable for a person in the defendant's position to become extremely angry.

This standard allows us to distinguish between the cases set forth in Part III. It does so by identifying the reason why considering the defendant's beliefs in cases such as the Aboriginal Man and the Devout Catholic seems appropriate, whereas considering the defendant's beliefs in cases such as the White Supremacist and the Conservative Pakistani Father seems inappropriate.

The White Supremacist defendant becomes extremely angry at seeing an inter-racial couple kissing in public. This defendant's extreme anger is properly blameworthy, as his anger does not manifest appropriate respect on his part for the interests of others. This is true even knowing that the defendant is a member (since childhood) of the Aryan Nation, and that his community has a shared commitment to white supremacy. The white supremacist beliefs promulgated by the Aryan Nation do not manifest appropriate respect for the interests of others—such as African-Americans (among others)—and thus the Aryan Nation

adherent's conduct in becoming extremely angry due to a violation of those beliefs is blameworthy.

Contrast this with the Aboriginal Man. Even though a non-Aboriginal person would not share the Aborigine's beliefs about "talking men's talk," a non-Aborigine might still be able to come to the conclusion that given everything we know about the Aboriginal defendant, his extreme anger at seeing another member of his tribe "talking men's talk" in front of uninitiated boys is not blameworthy. The Aboriginal Man's emotional response does not fail to manifest appropriate respect for the interests of others, even though the broader political community does not share his reasons for becoming angry.

Now consider the Devout Catholic. The Devout Catholic's belief that the church altar and the statues inside the church are holy and sacred is not necessarily a view shared or endorsed by the broader political community and is not a value affirmatively protected by state law.[155] At the same time, the Devout Catholic's belief in the sacredness of a statue of the Virgin Mary is not blameworthy or properly an object of condemnation. It does not manifest inappropriate respect for the interests of others. Accordingly, a jury evaluating the sufficiency of the provocation should consider the Devout Catholic's views that the church altar and statues were sacred and holy, even if the jurors themselves do not share those views.

This test does not dictate the outcome of each particular case. The jury must still determine whether it was "reasonable" for a Devout Catholic holding those views to become enraged at the sight of the desecration. If the case were extreme—such as the depiction in Roberto Bolaño's novel *2666*, in which a man known as "the Penitent" repeatedly urinates and defecates on Catholic church alters and decapitates many statutes, and also assaults priests—a jury might conclude that the provocation was sufficient and the Devout Catholic's extreme anger was reasonable. In contrast, if the Devout Catholic killed in a rage after seeing a tourist inadvertently sneeze on a holy statute, the jury would likely conclude that the provocation was not sufficient and the Devout Catholic's extreme anger was not reasonable. In any event, the "blameworthiness" standard properly frames the issue for the jury to make a judgment.

The Conservative Pakistani Father, like the Devout Catholic, relies on a mix of cultural and religious views. The Conservative Pakistani Father might believe (and might locate these beliefs in his cultural and religious community) that women

---

155. In some areas, such as heavily Catholic or Christian communities, these attributions of sacredness might indeed by broadly shared. In other areas they may not be. In either case, the sacredness of a religious shrine is not protected (as sacred) by state law. For example, a private citizen could construct or purchase a shrine to the Virgin Mary, and then treat that shrine disrespectfully—destroy it, or otherwise desecrate it. This conduct is not criminal, and indeed is a protected form of free speech. *See* Victor A. Kovner et al., *NYC Mayor Giuliani Loses Round One in Battle Against the First Amendment*, 17 COMM. LAW. 3, 3 (2000) (discussing the "preliminary injunction against the city and the mayor, thereby bringing a temporary halt to their efforts to punish the Brooklyn Museum for displaying art that the mayor had declared 'disgusting' and offensive to Roman Catholics").

should generally be submissive to men and that daughters should submit to the will of their fathers. The Conservative Pakistani Father is permitted to hold these beliefs. But the Conservative Pakistani Father's beliefs nonetheless run contrary to the state's commitment to sex and gender equality. It is blameworthy to become extremely angry in response to a daughter choosing to follow her own life path rather than submit to her father's will. Thus, a jury should reject the Conservative Pakistani Father's provocation claim. It might be true that he killed in a state of extreme anger, which made it much more difficult for him to control his conduct, and thereby renders him less culpable or blameworthy for his actions. But he is to blame for the reason he entered into that state of extreme rage in the first place, so he loses his claim to the jury's sympathy.

A similar analysis applies in *R. v. Humaid*, in which the defendant who killed his wife based on a suggestion of possible infidelity called an expert who "described Islamic culture as male dominated, preoccupied with the concept of 'family honour,' and particularly intolerant of female infidelity."[156] As the court explained, the problem in allowing Humaid to present this argument "is that the alleged beliefs which give the insult added gravity are premised on the notion that women are inferior to men and that violence against women is in some circumstances accepted, if not encouraged," and those beliefs "are antithetical to fundamental Canadian values, including gender equality."[157]

It is useful to contrast *Humaid* with the case of a Muslim man assaulted by a pigskin shoe. In each case, the defendant seeks to draw on particular religious and cultural sensitivities not shared by the broader political community, which greatly exaggerated the gravity of the provocation. For the man assaulted with the pigskin shoe, the religious and cultural taboo on products from pigs is not blameworthy, and thus can form a proper basis for his provocation claim. For Humaid, however, his extreme sensitivity to female marital infidelity, grounded in (claimed) religious and cultural beliefs, violates the political community's fundamental values—a commitment to gender equality. Thus, his provocation claim cannot properly be based on that blameworthy motivation.

C. Blameworthy Anger—Conceptual Foundations

The conceptual foundation of this "blameworthiness" test rests on the underlying justification for punishment in the criminal law—a normative judgment of moral blameworthiness. A defendant whose anger is an appropriate response to a provocation is not to blame for that anger.[158] This standard incorporates the insights of a number of commentators who likewise argue that the law must normatively evaluate the defendant's emotional response.

---

156. Way, *supra* note 52, at 7.
157. R v. Humaid, [2006] 81 O.R. 3d 456, 480 para. 93 (Can. Ont. C.A.); *see also* Way, *supra* note 52, at 11.
158. *See* Kahan & Nussbaum, *supra* note 22, at 306–07.

As Peter Westen explains, courts should not attempt to "incorporate" select traits or beliefs of the defendant onto an abstract "reasonable person."[159] Instead, courts should focus on the underlying reason we punish an individual in the criminal law. Courts should begin with the actual defendant and apply the appropriate moral filter to evaluate that particular defendant's reaction.[160] According to Westen, the legal system must determine whether the defendant's was "motivated in his conduct by proper regard for interests that the law seeks to safeguard," or whether the defendant "placed insufficient value on those interests."[161]

But rather than insist that the political community affirmatively share or endorse the defendant's anger, as Nourse and Westen do, the law should ask the more modest question of whether the community must affirmatively reject the defendant's anger—whether the reason for becoming angry is blameworthy. This allows the state to accommodate some degree of cultural and religious pluralism, while still articulating and protecting the core moral commitments of the broader political community.

Along similar lines, Andrew von Hirsch and Nils Jareborg note that "[w]hen one is wronged . . . anger may be the *appropriate* response."[162] This "appropriate anger" fits within the provocation doctrine as follows: "When one is angry, one's moral sense should serve only as . . . a bridle to the passions. . . . If the self-restraint fails, then one's moral inhibitions have not functioned as they should, and that hardly makes one any less to blame."[163]

In contrast, as von Hirsch and Jareborg state, in cases in which a defendant has been "wronged" and thus "is *properly* angry," the moral sense plays a different role.[164] "Far from having a purely suppressing role, one's sense of right and wrong is part of what prompts and gives legitimacy to the anger."[165] The defendant is less to blame "because the actor was moved to transgress in part because of, rather than despite, his sense of right and wrong."[166]

The defendant's ability to claim provocation thus rests in part on whether "there should be *good* reason for the actor's sense of injury or affront. When the actor has taken affront without just cause, then his sense of values is merely deficient."[167] In contrast, "[i]t is where an actor both *should* feel angry and yet must try to control that justifiable anger that he or she becomes less to blame for overstepping the

159. Westen, *supra* note 8, at 140, 148.

160. *Id.* at 140; *see also* Kahan & Nussbaum, *supra* note 22, at 306–07, 313–15 (making a similar claim under the rubric of an "evaluative conception" of emotion).

161. Westen, *supra* note 8, at 151.

162. Andrew von Hirsch & Nils Jareborg, *Provocation and Culpability*, *in* RESPONSIBILITY, CHARACTER AND THE EMOTIONS: NEW ESSAYS IN MORAL PSYCHOLOGY 241, 249 (Ferdinand Schoeman ed., 1988).

163. *Id.* at 250.

164. *Id.*

165. *Id.*

166. *Id.* at 251.

167. *Id.*

bounds."[168]

Kahan and Nussbaum make similar points, using their "evaluative conception" of emotion.[169] They state that the provocation doctrine should be available only when the defendant's "anger or rage is morally appropriate."[170] In contrast, the partial defense should not be available "in circumstances in which the actor's passion, however intense, reflects inappropriate valuations."[171] The provocation doctrine thus "condemn[s] acts that reflect at least some appropriate emotional motivations less severely than acts that reflect only inappropriate emotional motivations."[172]

None of these commentators, however, take the further step of explaining how to judge a defendant whose moral sense is motivated by a reason that the broader political community might not share. In cases like the Aboriginal Man as well as the White Supremacist, the defendant "was moved to transgress in part because of, rather than despite, his sense of right and wrong."[173] Yet the question remains whether that sense of right or wrong is correctly calibrated—whether "the actor's animus testifies to the soundness" or to the "deficiency . . . of his moral judgment."[174]

It is here that the "blameworthiness" standard helps clarify the judgment to be made. In the case of the Aboriginal Man, his extreme anger does not represent a "deficiency" in his moral judgment, because it is not blameworthy—even though it is not a moral judgment shared by the community. In the case of the White Supremacist, his moral judgment is "deficient" and thus his animus testifies to the blameworthiness of his moral judgment, not its soundness.

Westen proposes that blameworthiness be assessed by asking whether the defendant's "extreme emotional agitation . . . manifest[s] appropriate respect on his part for the interests of others?"[175] Westen mistakenly assumes that a defendant's extreme emotional agitation manifests appropriate respect for the interests of others only when the broader political community shares or affirmatively endorses the defendant's reason for becoming extremely angry.[176] But that is not the case. A defendant does not reveal himself to be blameworthy merely because the broader political community does not *share* his reasons for becoming extremely angry. A defendant can show appropriate respect for the interests of others

---

168. *Id.*

169. Kahan & Nussbaum, *supra* note 22, at 306–15.

170. *Id.* at 306–07.

171. *Id.* at 307.

172. *Id.* at 313.

173. von Hirsch & Jareborg, *supra* note 162, at 251.

174. *Id.*; *see also* Kahan & Nussbaum, *supra* note 22, at 313 (asking whether the defendant acted with "appropriate emotional motivations" or "inappropriate emotional motivations," without identifying a standard for making that judgment in contested cases).

175. *Id.* at 140.

176. *Id.*

even when the broader political community does not share or endorse the defendant's views. For example, an observant Muslim person or Jewish person who politely declines a host's offer of pork is not being rude, even if the host (like the broader political community) does not believe that pork is taboo.

Rather, a defendant reveals himself to be blameworthy when he is motivated by reasons that the broader political community condemns or affirmatively rejects.

D. Judging Emotion—A Role of the Liberal State?

Some may object to the idea of assessing the "blameworthiness" of a defendant's reasons for becoming angry, on the grounds that it seems problematic for the state to make a moral judgment about which types of personal motivations are "blameworthy" and which are not. In general, the proper function of the liberal state is to adjudicate harms inflicted by one individual against another, but not to engage in free-standing assessment (let alone punishment) of the blameworthiness of personal views held by individuals within the community.[177] Foundational liberal principles such as freedom of thought, freedom of speech, and freedom of conscience all presuppose the freedom of individuals within the broad political community to hold personal views at odds with others in the community, at odds with the majority, and at odds with the "official" views of the state.[178]

Kahan and Nussbaum recognize (and then reject) this objection. Under their "evaluative conception" of emotion, the state "appraises the evaluations internal to the offender's emotions as reasonable or not reasonable," and in doing so:

> takes a whole range of moral stands—about the importance of the murder of a child, about the (alleged) difference between the killing of an unfaithful wife's lover and the killing of an unfaithful girlfriend's lover, about the appropriateness or inappropriateness of the disgust occasioned by witnessing a homosexual act.[179]

They note that some might "think that it is improper for the law to get involved in such judgments at all in a liberal democratic society."[180] In a liberal state, individuals

---

177. JOHN STUART MILL, ON LIBERTY 68 (Penguin Books 1974) (1859) ("[T]he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant.").

178. *See* Everson v. Bd. of Ed., 330 U.S. 1, 15–16 (1947) ("The 'establishment of religion' clause of the First Amendment means at least this: Neither a state nor the Federal Government . . . can force nor influence a person . . . to profess a belief or disbelief in any religion. No person can be punished for entertaining or professing religious beliefs or disbeliefs."); Frazier v. Alexandre, 555 F.3d 1292, 1295-96 (11th Cir. 2009) ("Whether one calls it 'the right not to speak' or 'liberty of conscience,' the freedom to *think* and *hold beliefs* that do not comport with state orthodoxy is a basic part of this nation's foundation. . . .").

179. Kahan & Nussbaum, *supra* note 22, at 359.

180. *Id.*

> should be free to live their lives by their own conceptions of the good, deciding how much importance to accord to children, how much respect to the moral or religious standards that give rise to disgust at homosexual conduct, and so forth. Surely the law should be studiously neutral among competing conceptions of the good.[181]

Kahan and Nussbaum recognize that this normative judgment is both proper and inescapable. If the law chooses not to evaluate which types of anger are normatively blameworthy and which are not, then it is left to simply assess whether the defendant was in fact extremely angry (what Kahan and Nussbaum call the "mechanistic" conception of emotion).[182] Under that approach—refusing to normatively assess the defendant's anger—the law effectively concludes:

> that getting angry over the murder of one's own child and getting angry over the sight of two same-sex strangers making love differ in no salient way, that the only thing worth attending to here is the strength of the offender's anger, which might, of course, happen to be equally great in the two cases.[183]

By attempting to be "neutral" among differing conceptions of the good, the law would fail to reflect our intuitions and would "disregard what deeply rooted norms of reasonableness make central."[184]

In a provocation case, the state is not prosecuting an individual for "blameworthy" beliefs as such. Rather, the state prosecutes the individual for committing murder—intentionally killing another person, without justification. In response, the defendant seeks to mitigate his culpability for that wrongful killing. He does so by highlighting the fact that it was substantially more difficult for him to resist the urge to kill than is ordinarily the case, because he was in a state of extreme emotional disturbance. Due to that state, he pleads, he is substantially less culpable than one who intentionally kills in an ordinary state.

Yet the heat of passion doctrine implicitly recognizes that ordinarily, the act of becoming extremely angry is itself a blameworthy act. While it may be true that one who is extremely angry is less able to control his impulses, it is often true that the person is more culpable for being in a state of extreme anger. It is precisely this "loss of impulse control" feature of extreme anger—on which the defendant relies to plead mitigation—that makes it blameworthy to allow oneself to become extremely angry.

Even so, the heat of passion doctrine recognizes that there are some circumstances in which it is *not* unreasonable to become extremely angry—where the law should not expect what it ordinarily expects, namely that persons will refrain from entering that highly risky state of extreme anger (or at least, that they will retain

---

181. *Id.*
182. *Id.* at 360.
183. *Id.*
184. *Id.*

control over their violent impulses while in such a state). In fact, there are some cases in which "anger may be the *appropriate* response" to being wronged.[185]

In evaluating the defendant's reason for becoming extremely angry, the state is thus in the somewhat unique position of considering the defendant's plea that it was appropriate—or at least, not blameworthy—for him to become extremely angry in the circumstances. In this context, the objections stated above—about freedom of conscience and the proper role of the liberal state—lose much of their force. Rather than punishing people for unpopular thoughts, the state in a provocation claim is punishing the defendant for a wrongful killing. The state is willing to mitigate the defendant's crime as a concession to human weakness,[186] but not if the defendant's act of becoming extremely angry is itself blameworthy. If that is the case, the defendant has no proper claim on our sympathies.

Stephen Garvey argues that proposals to assess the moral rightness of the defendant's anger in voluntary manslaughter cases (this article's blameworthiness standard is one such proposal) are fundamentally illiberal.

> [I]f an actor finds himself saddled with an illiberal character, one composed of illiberal beliefs and desires, and if he should be unlucky enough to lose self-control and reveal, though not accept, his character's true colors in an otherwise criminal act, punishing him for having such a character is an illiberal use of state power.[187]

Garvey notes, "[t]he point here is *not* that a liberal state cannot hold its citizens in any way responsible for the content of their characters."[188] A liberal state is free to "educate its citizens in the ways of a liberal order" as well as "criticize and censure its illiberal citizens for their illiberal characters."[189] But the liberal state "*cannot* legitimately . . . heap *more* punishment on a citizen who, in the course of violating the criminal law, happens also to reveal the illiberal state of his soul."[190] Moreover, to the extent that

> one's beliefs are constituent elements of one's character over which one has no direct or immediate control, then the additional punishment imposed on the inadequately provoked actor, who neither accepts nor otherwise identifies with the offending belief, is punishment imposed for the content of his character, not for the character of his action.[191]

---

185. von Hirsch & Jareborg, *supra* note 162, at 249.

186. Joshua Dressler, *When "Heterosexual" Men Kill "Homosexual" Men: Reflections on Provocation Law, Sexual Advances, and the "Reasonable Man" Standard*, 85 J. CRIM. L. & CRIMINOLOGY 726, 747 (1995) (citations omitted).

187. Stephen P. Garvey, *Passion's Puzzle*, 90 IOWA L. REV. 1677, 1716 (2005).

188. *Id.* at 1717.

189. *Id.*

190. *Id.*

191. *Id.* at 1711.

Garvey's objection is unfounded. By denying the provocation partial defense to a defendant with blameworthy attributes or beliefs, the state is not punishing the person for those beliefs as such. Rather, the state is denying a mitigating partial defense to a person who has allowed those illiberal beliefs to motivate a violent outburst.

A person with a blameworthy character trait or belief is in possession of a latent risk—in a sense, this person is carrying around a bomb that might be triggered without warning and harm others. Consider a person with an extremely volatile temper. The standard metaphorical description of this person—they have a "short fuse"—maps precisely onto this "bomb" metaphor.

The "short fuse" metaphor suggests the following: For an ordinary person (not someone with a "short fuse"), an affront might ignite their "fuse," leading toward an eventual explosion. Because the fuse is not "short," this person has time to reflect on the affront, gain control of themselves, realize that the law prohibits them from reacting with violence, and seek vindication through other lawful mechanisms. This person has the time, given the length of their fuse, to snuff the flame and stop the explosion of violence.

For a person with a "short fuse," this opportunity for reflection and self-restraint is vanishingly short. After the affront ignites the fuse, the bomb of violence quickly explodes, before the cooler and more rational parts of the person's character have an opportunity to snuff the flame.

Focusing only on that moment, it seems improper to punish someone who explodes with violence due to their short fuse. If as Garvey posits, our characters are largely or perhaps entirely outside of our control,[192] then it seems improper to punish the man with a short fuse.

From a broader view, however, it is apparent that there is still adequate justification for blame. The man who has a (metaphorical) "short fuse" is like a person carrying around an explosive that can readily be triggered. Generally speaking, when persons transport explosives, our society imposes special obligations of care. For example, a trucking company transporting volatile chemicals is required to assume special burdens of care not required of those hauling an ordinary load.[193] The volatile chemical transporter might be required to use a specialized (and more expensive) trailer that minimizes the chance of explosion, to display special signs warning others of the risks, to take special precautions in areas with increased risks of explosion (such as a gas station), and to avoid entirely certain areas with increased risks (for example tunnels).[194]

---

192. *Id.*

193. *See, e.g.*, Hazardous Materials Transportation Act of 1975, Pub. L. 93-633, 88 Stat. 2156 (1975) (codified as amended at 49 U.S.C. §§ 5101–28 (2012)).

194. *See* Shippers—General Requirements for Shipments and Packagings, Hazardous Materials Regulations, 49 C.F.R. Part 173 (2017).

So too with our short-fused man. We might reasonably expect this person, so far as possible, to cultivate habits of patience and self-control. We might reasonably expect him to avoid situations in which he knows his short fuse is particularly likely to be lit: drinking at rowdy bars, or hanging around his ex-girlfriend who often taunts him. We might expect him to seek assistance from his friends and family for help in controlling his short fuse, including intervening if necessary when he is getting angry.

Accordingly, it is just to punish the short-fused man as a murderer—to deny him the partial defense of provocation—when he flies into a rage and purposely kills in response to a provocation that would not similarly enrage a reasonable (not short-fused) person. The short-fused man is not being punished for having a short fuse. He is being denied a mitigating defense for failing to be aware of the considerable risk his character posed to others—the short-fused bomb he was carrying around—or, if he was aware of it, failing to take the necessary special precautions to prevent his short-fused bomb from exploding and killing an innocent person.

Now turn to a person with a blameworthy belief (or in Garvey's term, an illiberal belief), such as a white supremacist vehemently opposed to inter-racial romance. This person is free, in a liberal state, to hold this repugnant view. He has some degree of freedom to make his own life choices according to that repugnant view—avoiding intimate partners of other races, and urging his friends and family to do likewise. At the same time, he must understand that his views run contrary to the fundamental values of the state and polity in which he lives. His freedom to live according to his white supremacist views has limits imposed by the state, based on the impacts those beliefs may have on others. If he is an employer or rents out housing to the public, for example, he is prohibited from acting on his white supremacist beliefs in hiring or choosing renters.[195]

Moreover, the white supremacist—like the short-fused man—must understand that, in a sense, he carries around with him an explosive device. He is liable to become extremely angry under certain circumstances in which, he must understand, he will receive no sympathy from the state. He knows that attending a public park may well expose him to the sight of inter-racial couples showing affection. He is free, in the liberal state, to harbor his hatred of this sight. He even can vent verbally, casting offensive aspersions on the couple, so long as his verbal comments do not constitute prohibited threats or harassment.

However, he must take great care—if he wishes to avoid the sanction of the criminal law—that his illiberal beliefs do not manifest themselves in action, even in a fit of anger. Garvey notes that if such a person purposefully channels his illiberal beliefs into criminal action—such as going to a park looking for a black

195. *See* Civil Rights Act of 1964, tit. 7, 42 U.S.C. §§ 2000e-1 to -17 (2012).

man to assault—he can properly be punished for a hate crime.[196] He concedes that this is not a punishment of the illiberal belief as such, but a punishment for channeling that belief into violent, criminal action.[197]

A similar claim can be made in the provocation context. The white supremacist must take care that his illiberal beliefs—the sensitive explosives he has chosen to carry around with him—do not become momentarily enflamed by some provocation (an inter-racial couple kissing in the park) in a way that causes him to "lose control" and kill.

It is appropriate for the liberal state to place this special obligation of self-control on him, vis-à-vis his blameworthy or illiberal beliefs. It is appropriate for the liberal state to deny his plea for mitigation when he succumbs to the great anger arising suddenly from the illiberal bomb (his white supremacist beliefs) he has chosen to carry around.

Viewed in this way, the state is not punishing the white supremacist for his white supremacist beliefs as such, just as it does not do so in a hate crimes prosecution. Rather, in both cases the state punishes the white supremacist for channeling his illiberal beliefs into criminal action. In the case of a hate crime, he channels that belief into action deliberately and purposefully. In the case of the provocation doctrine, he fails to maintain adequate safeguard so that his illiberal beliefs do not trigger the extreme anger for which he should know the state will offer him no sympathy.

Persons with (partly) illiberal characters have an obligation to take care that their illiberal views—views they are free to hold, but which they understand contravene the core values of the state in which they reside—are not channeled into violent outbursts. If that occurs, they cannot properly expect mitigating sympathy from the state, notwithstanding the sincerity of their anger and their genuine reduction in self-control.

As noted at the outset, substantial operational questions remain as to whether the legislature, the judge, the jury, or some combination of those institutions, is best suited to determine which types of motivations are "blameworthy" and which are not. Cynthia Lee argues at length that juries should be allowed to make these normative determinations, albeit with instructions from the judge not to rely on sexism, racism, homophobia, transphobia, or similar disfavored reasons.[198] I do not purport to resolve these important operational concerns here. My focus, instead, is on identifying and articulating the correct normative standard, regardless of which legal institution implements and applies that standard.

---

196. Garvey, *supra* note 187, at 1711–13.

197. *Id.* at 1713.

198. *See* LEE, *supra* note 9, at 217–25, 252–53, 258–59; Lee, *supra* note 8, at 536, 564–66; Lee & Kwan, *supra* note 20, at 119–32; *see also* Dressler, *supra* note 9, at 1001.

VI. "GAY PANIC," "TRANS PANIC," AND POSSESSIVE MEN KILLING INDEPENDENT WOMEN

Defendants seek a provocation jury instruction in several recurring—and highly controversial—contexts: men killing their intimate female partners out of claimed jealousy, men killing other men in response to an unwanted sexual advance (so-called "gay panic" cases), and men killing a trans woman upon discovering that she is biologically male (so-called "trans panic" cases). These defendants want the jury to consider the claimed provocation from the defendant's point of view. In each of these contexts, commentators have criticized the provocation doctrine for enabling or excusing sexism, homophobia, and transphobia.[199] Some have gone so far as to call for the provocation doctrine to be abolished due to these concerns.[200] Other commentators, including many who share these concerns, nonetheless defend the provocation doctrine and seek to retain it in at least some cases.[201]

The "blameworthiness" principle articulated above provides the correct normative framework for mediating these disputes. At a deep level, provocation claims in these contexts share a fundamental similarity with provocation claims based on minority cultural or religious views. Whether the defendant comes from our own culture (or a sub-culture within it), or from a wholly foreign culture, the basic challenge is the same: should the jury view the sufficiency of the provocation from the defendant's perspective, or should the jury hold the defendant to some objective standard?

The possessive man who kills his wife when she attempts to leave him does not necessarily belong to an identifiable "cultural minority" (although he might). This defendant is asking the jury to view the reasonableness of the provocation from "his perspective"—a perspective internal to his view of the world.[202]

Cynthia Lee has argued throughout her work on the provocation doctrine that contested normative judgments, such as those involving issues of race, gender, and sexual orientation, should be brought to the forefront and made salient to jurors, rather than being permitted to silently or subconsciously influence decision-making.[203]

---

199. *See infra* notes 204, 269, 275, and 285.

200. *See, e.g.*, JEREMY HORDER, PROVOCATION AND RESPONSIBILITY 193–97 (1992).

201. *See, e.g.*, LEE, *supra* note 9, at 203–75; Dressler, *supra* note 9, at 994–95; Nourse, *supra* note 17, at 1389 ("My proposal seeks to reconstruct, rather than abolish, the defense.").

202. *See* LEE, *supra* note 9, at 110–11 ("Contrary to popular perception, immigrant and racial minority defendants are not the only ones who try to mitigate their criminal charges by relying on cultural norms. Every time a White heterosexual male murder defendant argues he was reasonably provoked into a heat of passion by his female partner's infidelity or by a gay man's sexual advance, he *seeks* to mitigate his charges by relying on cultural norms.") (emphasis added); Ben-David, *supra* note 9, at 10 ("One of the criticisms made against recognizing cultural arguments as part of the provocation doctrine is that it improves the situation of men who killed their spouses out of jealousy.").

203. *See* LEE, *supra* note 9, at 203–75; Lee, *supra* note 8.

The "blameworthiness" principle dovetails with Lee's approach, by helping articulate for jurors (or other decision-makers) the nature of the normative judgment to be made. In some provocation cases, jurors may feel conflicting intuitions about whether a given defendant encountered a provocation that was "adequate" or "sufficient," because some of the defendant's reasons for becoming extremely angry were appropriate—not blameworthy—while other parts of the defendant's reasons were blameworthy. The "blameworthiness" principle helps judges, lawyers, and jurors think clearly about these contested normative issues, by framing the correct standard against which to judge individual defendants.

A. "Blameworthiness" and the Feminist Critique of Provocation

Feminist scholars have criticized provocation law as a doctrine that primarily benefits men, reinforces cultural stereotypes of masculinity and gender hierarchy, and wrongly mitigates punishment for men who are highly violent and pose a serious risk of danger to women in the future.[204] At the same time, many commentators, including some of these feminist critics, believe that the provocation doctrine (properly understood) serves important values and should not be abolished.[205]

In the traditional legal imagination, a paradigmatic heat of passion case involves sexual infidelity: a man discovering his wife in the act of having sex with another man and killing either the wife or the other man (or both).[206] Victoria Nourse's research revealed that "the conventional image of passionate homicide—sexual betrayal, love triangles, sordid affairs—represents a flawed view."[207] In many cases, possessive men kill intimate partners not due to sexual infidelity, but rather in the context of the woman leaving (or attempting to leave) the relationship.[208]

Donna Coker has likewise questioned the traditional assumptions surrounding violent male responses to female infidelity—or, perhaps, female independence. Coker first questions "the belief that violence in response to a wife's provocation—in this context, the wife's adulterous conduct—is an uncontrollable response, which in turn reinforces the belief that intervention can have little

---

204. *See, e.g.*, HORDER, *supra* note 200 (arguing that the doctrine should be abolished because it is biased toward men who kill women); LEE, *supra* note 9, at 20–33; Coker, *supra* note 18, at 75–76; Mahoney, *supra* note 18; Elizabeth Rapaport, Comment, *Capital Murder and the Domestic Discount: A Study of Capital Domestic Murder in the Post-* Furman *Era*, 49 SMU L. REV. 1507, 1546 (1996); Laurie J. Taylor, *Provoked Reason in Men and Women: Heat-of-Passion Manslaughter and Imperfect Self-Defense*, 33 UCLA L. REV. 1679 (1986).

205. *See, e.g.*, *supra* note 201.

206. *See* LEE, *supra* note 9, at 17–25.

207. Nourse, *supra* note 17, at 1344.

208. *Id.* at 1343–58; *see also* Coker, *supra* note 18, at 76 (criticizing the legal system for failing to recognize that abusive men who kill their intimate partners are often not acting in an uncontrolled, provoked heat of passion, but in a predictable pattern of escalating violence that is likely to recur); Mahoney, *supra* note 18, at 78–79 (arguing that the law often improperly casts acts of separation by women as acts of infidelity or provocation).

deterrence or prevention impact."[209] She argues instead that many men who kill their intimate partners have a long history of violence against women, and thus their act of killing is less an uncontrollable emotional response, and more of a predictable manifestation of a grave character flaw—a disregard for women's rights and interests.[210]

Relatedly, Coker questions the notion that "the wife-killer who kills in response to his wife's 'provocative' conduct is seen as an unlikely recidivist and therefore less dangerous."[211] Coker argues that in many cases, the defendant's history shows he is generally violent toward women and does indeed pose a serious and predictable threat of future violence toward women.[212]

Even so, many commentators defend the provocation doctrine at least in some circumstances. Several feminist critics recognize that provocation can appropriately mitigate the culpability of some women who kill their abusive partners, even when those killings fall short of self-defense.[213] Moreover, there is no doubt that a sexual betrayal by a spouse or long-time intimate partner can inflict profound emotional and psychological trauma. In at least some cases, a response of extreme anger may be appropriate in response to sexual infidelity.

Feminists have rightly pointed to cases in which the provocation doctrine has wrongly served to partly excuse a defendant who has no rightful claim on our sympathies. At the same time, there are other cases involving infidelity in which the defendant's extreme emotional response seems more appropriate, in which the provocation claim should be available.[214] The right normative question is whether the defendant's extreme emotional response was blameworthy—whether the reasons the defendant became extremely angry reflected blameworthy traits or beliefs, such as gender hierarchy.

Consider the infamous case of *People v. Berry*,[215] criticized by many as showing the provocation doctrine at its worst.[216] *Berry* held that the twenty-hour period between the victim's alleged provocation and the defendant's act of killing was not too long to merit a provocation instruction.[217] Notwithstanding this relatively lengthy period, and the rule that provocation is unavailable if there was an adequate "cooling off" period, the court concluded that "the long course of provocatory conduct [by the victim], which had resulted in intermittent outbreaks of rage under specific provocation in the past, reached its final culmination in the

209. Coker, *supra* note 18, at 76.
210. *Id.* at 90–94, 128–29.
211. *Id.* at 76.
212. *Id.* at 90–94, 128–29.
213. *See* LEE, *supra* note 9, at 252; Nourse, *supra* note 17, at 1366–67.
214. An example of one such case, discussed below, is People v. Le, 69 Cal. Rptr. 3d 831 (Cal. Ct. App. 2007).
215. 556 P.2d 777 (Cal. 1976).
216. LEE, *supra* note 9, at 17–19; Coker, *supra* note 18, at 118–28; Mahoney, *supra* note 18, at 78–79.
217. 556 P.2d at 781.

apartment when [she] began screaming."[218]

In the case, Albert Berry, age 46, married the much younger Rachel Passah, a 20-year old Israeli woman.[219] Three days after the wedding, Rachel left and traveled alone back to Israel, returning to California after about six weeks.[220] Albert testified that upon her return, Rachel informed him she had fallen in love with, and slept with, a man in Israel (named Yako).[221] According to Albert, what followed was "a tormenting two weeks in which Rachel alternately taunted defendant with her involvement with Yako and at the same time sexually excited defendant, indicating her desire to remain with him."[222]

Ten days after her return from Israel, Albert choked Rachel so severely that she became unconscious.[223] She was treated at a hospital and reported the attack to San Francisco police. Soon after, a warrant was issued for Albert's arrest.[224] While Rachel was in the hospital, Albert took his clothes and moved out of their apartment.[225]

Within a day or two, Albert went back to the apartment to talk with Rachel, but she was not present.[226] He slept at the apartment, waiting for her, and she returned the next morning around 11 a.m.[227] When Rachel returned and saw Albert, she said, "I suppose you have come here to kill me." Albert responded "yes," then "no," and then "yes" again.[228] Rachel began screaming, they struggled, and Albert "finally . . . strangled her with a telephone cord."[229]

The California Supreme Court reversed Berry's conviction for murder in the first degree, accepting Berry's argument that he was entitled to a provocation instruction based on these facts.[230]

Donna Coker argues that the court allowed the provocation defense to be used to mitigate a killing based on factors that, in reality, showed the defendant to be more culpable, not less culpable.[231] Berry's defense expert in essence blamed the victim—Rachel—for the blameworthy attitudes and misconduct of the defendant. Berry's past "propensity to assault wives and lovers under circumstances in which he claimed the woman's infidelity provoked him" was used to claim that Rachel

---

218. *Id.*
219. *Id.* at 778.
220. *Id.*
221. *Id.* at 779.
222. *Id.*
223. *Id.*
224. *Id.* at 778–79.
225. *Id.* at 779.
226. *Id.*
227. *Id.*
228. *Id.*
229. *Id.*
230. *Id.* at 778, 782.
231. Coker, *supra* note 18, at 118–28.

abused and exploited a "weakness" of Berry's.[232] "The defense portrayed Rachel as a vindictive—albeit confused—woman who sexually used and abused Berry in order to gain her own death."[233]

In reality, Coker points out, "the truth is that Berry didn't kill Rachel until it appeared that she might make good on her threat to leave him."[234]

In *Berry*, the court failed to properly assess whether Berry's reasons for becoming extremely angry at Rachel manifested an appropriate respect for the interests of others. As Coker and others have explained, Berry's anger stemmed in large part from Rachel's decision to leave him, and his improper—blameworthy—impulse to control her.

Victoria Nourse likewise highlights a series of cases in which juries were instructed on provocation or "extreme emotional disturbance," even though the facts showed that the defendant's anger and rage reflected his desire to control his current or former partner, and a refusal to allow her to choose to leave him.[235] To the extent the facts show that a defendant's rage is the result of beliefs in gender hierarchy and female submission, he should be denied an instruction on provocation because those beliefs fail to manifest appropriate respect for the interests of others—the freedom of women to choose their intimate associations and living arrangements.

In some other cases falling into the same general category—men claiming provocation due to female infidelity—the facts paint a very different picture, and show a legitimate grounds for allowing provocation claims because they do not rest on objectionable norms of male control and female submission. For example, in *People v. Le*, defendant Johnny Viet Le killed Van Truoung (known as Thang) who had been having an affair with Le's wife, Nga Pham.[236] Le and Pham met and were married in Vietnam, and came to the United States in 1989.[237] They had three children together, and Pham testified that Le "was a 'very good husband,' that he loved her 'a lot,' and that 'after 21 years of marriage, he never raise his hand on me.'"[238]

Pham began having an affair with Thang around 2002. Le discovered the affair in November, but the couple was able to work things out. Pham testified she promised Le she would stop the affair, and Le cut off a portion of one of his fingers as a symbol of forgiveness and starting anew.[239]

---

232. *Id.* at 120–24.

233. *Id.* at 122.

234. *Id.* at 128; *see also* Mahoney, *supra* note 18, at 78–79.

235. Nourse, *supra* note 17, at 1342–43 (citing State v. Utz, 513 A.2d 1191, 1192 (Conn. 1986); Smith v. Commonwealth, 734 S.W.2d 437, 440 (Ky. 1987); State v. Wille, 858 P.2d 128, 130 (Or. 1993)).

236. People v. Le, 69 Cal. Rptr. 3d 831, 833–35 (Cal. Ct. App. 2007).

237. *Id.* at 833.

238. *Id.*

239. *Id.*

Notwithstanding these events, the affair continued, and Pham loaned Thang $10,000.[240] Le again learned about the affair, and began monitoring Pham's phone calls and movements.[241] In 2003, Le and Pham traveled (with their children) to Vietnam, where Le told Pham's parents about the affair.[242] Pham's parents implored her to stop seeing Thang, and Pham again promised to terminate the affair.[243]

Le again learned of the affair, and shaved his head and considered becoming a monk. He asked Pham for a divorce, but she refused. In May 2003, they decided to move the family to Texas.[244] During the drive, Thang called Pham on her cell phone. She put the phone on speakerphone so Le (and another relative in the car) could hear. Thang asked for $10,000 for a trip to Vietnam and for money to start a business, and also asked if she had life insurance.[245]

In mid-May, Le and Pham returned to California to get their children, who had stayed behind to finish school. Le told Pham he wanted to contact Thang and meet with him, in the hopes of convincing Thang to terminate his relationship with Pham. Pham did not want Le to meet with Thang, and Le became extremely angry, believing that Pham still loved Thang.[246]

Le went for a bicycle ride and bought a meat cleaver. He set up a meeting with Thang, and then attacked and killed him with the cleaver.[247] Soon after, Le turned himself in to the police.[248]

The defense called an expert in "cross-cultural psychiatry," Paul Leung. Leung "described the differences between acute and chronic stress and the importance of 'saving face' in Asian culture."[249] He testified about how long-term, chronic stress could "bring a person to act rationally while in the process of committing an irrational act."[250]

The defense objected to a jury instruction stating that "mere words" can never constitute adequate provocation. The jury convicted Le of second-degree murder. On appeal, the California Court of Appeal reversed, holding that this instruction was erroneous and that the evidence in the case supported an instruction on voluntary manslaughter under the provocation doctrine.[251]

---

240. *Id.*
241. *Id.*
242. *Id.*
243. *Id.*
244. *Id.*
245. *Id.* at 834.
246. *Id.*
247. *Id.* at 835.
248. *Id.*
249. *Id.* at 836.
250. *Id.*
251. *Id.* at 832, 836, 839–40.

For present purposes, the case illustrates how the blameworthiness principle serves to distinguish cases in which the provocation doctrine should not be available from cases in which it should.

In one sense, Le's case fits into the long line of cases in which men commit a homicide in response to infidelity by a wife or intimate partner (in this case, Le killed his wife's lover rather than his wife). As critics have argued, the men in many of these types of cases should not receive a provocation instruction.[252] But Le's case, in my view, clearly merits a provocation instruction.

Unlike many such cases, Le's extreme anger related to his wife's infidelity does not seem to reflect blameworthy beliefs or personality traits. The evidence did not show that Le was a jealous and possessive man. He did not abuse his wife, seek to control her, or seek to prevent her from ending her relationship with him. On the contrary, he repeatedly sought to reconcile with her. He traveled with her to Vietnam to try to enlist her family's help in mending their relationship. He took the extraordinary step of cutting off part of one of his own fingers as a serious symbol of forgiveness and reconciliation. When problems persisted after all of these steps, he sought to end their relationship with a divorce, which she resisted. There was no evidence he had a volatile temper.

Moreover, the cultural testimony in Le's case did not ask the jury to accept or endorse blameworthy cultural norms, as in some other cases. In *R. v. Humaid*, the defendant in essence wanted the jury to view the provocation from the perspective of a "reasonable conservative Islamic man," in particular (according to the defendant) a reasonable man from a culture that was "male dominated, preoccupied with the concept of 'family honour,' and particularly intolerant of female infidelity."[253]

In contrast, Le's expert did not seek to invoke patriarchal cultural norms to justify Le's anger. Rather, he described the norm of "saving face" in Asian cultures, and described how a person might seem to be acting rationally while still being in the midst of extreme emotion. These cultural or psychological features—assuming the jury is persuaded that they exist—are not blameworthy. The Asian cultural norms regarding "saving face" may be different than prevailing norms in the United States. They may not be affirmatively shared by most Americans. Yet norms of "saving face" do not seem affirmatively blameworthy—they do not seem to place insufficient weight on the legitimate interests of others.

Some of the most difficult cases—both for our intuitions and for a clear conceptual framework—involve circumstances in which the defendant is partly but not entirely blameworthy for becoming extremely angry. In these cases, the

252. *See supra* note 204.

253. Way, *supra* note 52, at 7 (quoting R. v. Humaid, [2006] 81 O.R. 3d 456, 473–74 para. 67 (Can. Ont. C.A.)).

"blameworthiness" test identifies the proper normative standard that lawyers, courts, and juries should use to evaluate competing claims.

One notable case involving partial blameworthiness is *Bedder v. Director of Public Prosecutions*,[254] discussed by many commentators.[255] Bedder was an eighteen-year-old man who knew he was impotent.[256] He went to a prostitute to have sex, and predictably failed to perform. The prostitute jeered at him for his failure, and tried to leave. Bedder held on to her; she hit and kicked him in response. He then stabbed her to death.[257]

Victoria Nourse notes that "[t]raditionally, scholars have had great sympathy for the youthful defendant but have had trouble defending the proposition that impotence is a characteristic of a reasonable person."[258] Feminist scholars, in contrast, have "seen this sympathy as another example of the law bowing to male concerns about virility."[259]

Nourse evaluates the case under her proposed standard—namely, whether "the defendant appeals to the very emotions to which the state appeals to rationalize its own use of violence."[260] Applying that rule, Nourse argues that "because the defendant's virility was a matter of particular sensitivity to him" (rather than a concern shared by the broader community, as reflected in the law), the "jury should have been instructed that it could not return a manslaughter verdict."[261]

At the same time, Nourse concedes that "there is a possible claim for compassion in Bedder," grounded in the "unstated analogy between Bedder and those who are victimized for their attributes; that is, because of their handicap, race, or sex."[262] On the other hand, even if Bedder's case is viewed as a "handicap," Bedder might still be faulted for "blaming the victim for his own handicap."[263]

Nourse concludes that framing the issue this way "does not require us to ask whether impotence is a characteristic of a reasonable person, nor does it ignore the potential for biased generalizations (of men or women). It exposes the judgments so that a choice may be made."[264]

The "blameworthiness" standard guides this normative judgment, as well as helping to isolate and separate the issues for the parties and the fact-finder. As Nourse recognizes, Bedder's reason for becoming extremely angry can be characterized in a variety of ways:

---

254. [1954] 2 All ER 801, 1 WLR 1119 (UK).

255. *See, e.g.*, Nourse, *supra* note 17, at 1402–03; Westen, *supra* note 8, at 149–50; Meir Dan-Cohen, *Responsibility and the Boundaries of the Self*, 105 HARV. L. REV. 959, 993 (1992).

256. *Bedder*, 2 All E.R. at 801–02.

257. *Id.* at 802.

258. Nourse, *supra* note 17, at 1402.

259. *Id.*

260. *Id.* at 1338.

261. *Id.* at 1402.

262. *Id.*

263. *Id.*

264. *Id.* at 1403.

1. Male Chauvinism: Bedder became enraged because of a male chauvinism that found it impertinent for a woman (who should be submissive) to mock the sexual performance of a man (who should be dominant).
2. Physical Disability: Bedder became enraged because the victim made fun of him for a physical disability over which he had no control.
3. Sexual Hubris: Bedder became enraged because he felt humiliation at failing to perform sexually, and he "took out" his humiliation on the victim even though she was not to blame for his failure and he had knowingly sought out the encounter that led to his predictable humiliation.
4. Failure to be Masculine: Bedder became enraged because the victim made fun of him for failing to meet her expectations of "stereotypical sexual performance standards for adult males."[265]

Of course, it could be the case that some or all of these explanations are present at the same time. Perhaps the victim invoked sexual stereotypes (Failure to be Masculine), but that taunting was particularly effective against Bedder because he himself already had a strongly held belief in sexual stereotypes (Male Chauvinism). It is in this way that we see *Bedder* as a possible example of "partially blameworthy" motivations. It could be the case that Bedder's motivations were partly blameworthy, and partly not blameworthy.

In part the dispute in *Bedder* is factual—what reason(s) really prompted Bedder's extreme anger? In a case before a jury, the prosecutor and the defense lawyer might well try to characterize Bedder's extreme anger in some of the different ways suggested above, or might come up with additional and perhaps more persuasive interpretations of what prompted his anger. The "blameworthiness" standard frames the arguments a prosecutor or a defense lawyer would make; it serves as the benchmark against which the jury must evaluate those competing characterizations.

After the factual questions are resolved, the normative question remains—were those reasons blameworthy?

If Bedder's anger stems from a belief that woman should be sexually submissive and males sexually dominant (Male Chauvinism)—and the victim violated those norms by taunting him over a matter of male sexual performance—then his anger stems from a blameworthy attitude of male sexual dominance over females.

If Bedder's anger stems from personal humiliation that he "took out" on the victim (Sexual Hubris), that anger is likewise blameworthy. His sense of humiliation is not blameworthy, but his emotional response to translate humiliation into anger at a person who did not wrongfully cause that humiliation is blameworthy. That is particularly true in a case like this, in which Bedder himself chose the encounter he knew would risk his own humiliation.

---

265. *Id.* at 1402.

At the same time, Bedder is not at fault for being impotent. If his anger stems from being mocked for a physical disability (Physical Disability), that anger is not a blameworthy emotional response—just as a person who is blind is not blameworthy if she becomes angry at being mocked for her lack of sight.

If Bedder's anger was prompted by the prostitute's invocation of stereotypical male sexual performance standards (Failure to be Masculine), his anger is likewise not blameworthy. There is a very fine line between this way of characterizing things and the first reason, "Male Chauvinism." In the first case, it is Bedder himself who holds sexist stereotypes about male and female sexual roles and performance, and it is his own sexist stereotypes that provoke his anger. In this case, however, it is not Bedder but the victim who invokes blameworthy sexual stereotypes.

In the end, it seems that Bedder's anger is partly defensible as an appropriate (non-blameworthy) response to being taunted for a physical disability, and also is partly blameworthy. The extreme nature of Bedder's anger seems hard to explain based only on the taunting for a physical disability. What more plausibly led to Bedder's *extreme* anger was his blameworthy fury at being humiliated—a humiliation he brought upon himself—as well as the affront to his masculinity by a woman. If those blameworthy motivations are removed from the situation, it does not seem likely that the provocation was sufficient to incite a reasonable person into extreme anger.

Feminists have made convincing claims that the provocation doctrine has, too often, wrongly been available to mitigate punishment for men who are highly violent and pose a serious risk of danger to women in the future—men whose anger stems from inappropriate cultural stereotypes of masculinity and gender hierarchy.[266] Judges and juries should reject provocation claims to the extent the defendant's reason for becoming extremely angry is blameworthy—for example, when it stems from cultural stereotypes of masculinity and gender hierarchy. At the same time, some cases do not reflect those blameworthy rationales, and in those cases the provocation doctrine should be available.

B. Gay Panic, Trans Panic, and Blameworthiness

Another controversial and much-criticized use of the provocation doctrine is in the context of so-called "gay panic" or "trans panic" cases.[267] Joshua Dressler refers to "gay panic" cases with the more neutral formulation of "nonviolent homosexual advance," or NHA, cases.[268]

---

266. *See supra* note 204.

267. LEE, *supra* note 9, 67–95; Lee, *supra* note 8, at 564; Lee & Kwan, *supra* note 20; Mison, *supra* note 20; Strader et al., *supra* note 56, at 1501–07.

268. Dressler, *supra* note 186, at 727.

Robert Mison argues that a nonviolent sexual advance by a person of the same sex should never be viewed as "sufficient provocation."[269] Permitting the provocation claim in a "gay panic" case, he argues, is "a misguided application of provocation theory and a judicial institutionalization of homophobia."[270] He argues that a correct application of the "reasonable person" standard would dictate this result: "a reasonable person should not be provoked to kill by such an advance. The reasonable man should not possess prejudices and biases such as homophobia and heterosexism."[271] He therefore concludes that judges should "find as a matter of law that a homosexual advance is insufficient provocation," and thus prohibit juries from considering this claim.[272]

Similar arguments have been made in so-called "trans panic" cases, in which the defendant claims "that his violent acts were triggered by the revelation that another person, sometimes with whom he has been sexually involved, is transgendered."[273]

Joshua Dressler responds that Mison's categorical rule is misplaced. Dressler argues that in some cases in which men are provoked during a sexual advance from another man, a manslaughter instruction may be appropriate.[274] Dressler agrees with Mison in part. For one, Dressler claims that "the Reasonable Man in NHA cases is not homophobic."[275]

At the same time, Dressler argues that "[i]t is not necessarily the case that a person who kills after a NHA does so as the result of intolerance, bigotry, or homophobia."[276] As a general matter, "an unwanted sexual advance is a basis for justifiable indignation."[277] The appropriateness of the response—how extreme the anger is—depends on the nature of the sexual advance. For example, "[i]f a victim lightly touches a defendant on the shoulder and asks, 'Do you want to have sex with me?,' such a solicitation should not result in an instruction on manslaughter."[278] In contrast, Dressler argues that "[m]ost men—including non-homophobic heterosexuals *and gay men*—would justifiably become indignant if a stranger nonconsensually touched their genitals, fondled their buttocks, or committed a sexual act upon them while they slept."[279] Dressler concludes that judges should instruct juries to evaluate whether legitimate indignation caused the defendant's extreme emotional response, while trying to prevent juries from returning a

269. Mison, *supra* note 20, at 135–36.
270. *Id.* at 136.
271. *Id.* at 161.
272. *Id.* at 176–77.
273. Steinberg, *supra* note 20, at 501.
274. Dressler, *supra* note 186, at 753–56.
275. *Id.* at 752.
276. *Id.* at 754.
277. *Id.*
278. *Id.* at 760.
279. *Id.* at 754.

manslaughter verdict based on homophobia.[280]

Cynthia Lee likewise prefers for such normative questions to be aired before the jury, along with judicial instructions that jurors "should try not to let" norms such as "masculinity norms, heterosexuality norms, and race norms . . . bias their decision making."[281]

Notwithstanding some disagreements, it is notable that Mison, Dressler, and Lee all agree that if a defendant killed in response to a non-violent homosexual advance because of the defendant's homophobia, or prejudice toward homosexuality, that defendant should be convicted of murder, not the lesser charge of manslaughter.[282]

A general conceptual explanation of this point of agreement is that acting out of fear, prejudice, or hatred toward homosexuality is blameworthy. In contrast, several of the other possible motivations that Dressler mentions are not blameworthy. It is appropriate to respond with indignation to an unwanted sexual advance. If that sexual advance involves a physical violation of one's body, even stronger indignation is warranted. In a sufficiently egregious case, it may be appropriate to respond with extreme anger.

Of course, in most cases it seems hard to explain a response of extreme anger to a nonviolent homosexual advance without reference to a deep-seated hatred or fear of homosexuality. If the defendant's blameworthy homophobia is the key factor that turned what might have been moderate indignation into extreme, violent rage, then it is wrong to mitigate the defendant's act of killing from murder to manslaughter. The defendant lost his temper, and would not have lost it—not in the extreme way that he did—had he "manifest[ed] appropriate respect . . . for the interests of others."[283]

The blameworthiness test should also be applied in so-called "trans panic" cases, in which a defendant kills an intimate partner upon discovering that the partner is transgender.[284] To date, most commentators agree that certain types of motivations in these cases are normatively blameworthy, and thus should not form a basis for adequate provocation: claims that the defendant responded with anger due to a hatred and disgust for transgender persons, or took offense at a violation of traditional gender roles.[285]

---

280. *Id.* at 758–62.

281. LEE, *supra* note 9, at 252–53.

282. *See id.*; Dressler, *supra* note 186, at 758–62; Mison, *supra* note 20, at 172.

283. Westen, *supra* note 8, at 140.

284. *See* Bradford Bigler, Comment, *Sexually Provoked: Recognizing Sexual Misrepresentation as Adequate Provocation*, 53 UCLA L. REV. 783, 827–28 (2006); Lee & Kwan, *supra* note 20, at 77; Steinberg, *supra* note 20, at 500–02, 509–10.

285. *See* Lee & Kwan, *supra* note 20, at 126 ("A prosecutor in a trans panic case can similarly argue that, if the jury finds the defendant's anger was the product of unacceptable transphobia that the defendant should have controlled, they should fault him for his anger and return a verdict of murder."); *id.* at 113 ("[T]he law of

Some, however, have argued that there may be adequate provocation in some cases on a theory of sexual deception: that it is reasonable to react with anger (even extreme anger) upon discovering that an intimate partner, believed to be a woman, is biologically a male.[286] Bradford Bigler argues that this type of sexual deception can reasonably give rise to extreme emotion, due to justified feelings of betrayal of one's sexual autonomy.[287]

Other commentators disagree. Victoria Steinberg argues that "revealing one's biological sex does not constitute sufficient provocation."[288] Cynthia Lee and Peter Kwan likewise conclude that "it is not normatively appropriate to be outraged by the discovery that one's intimate partner is a transgender person."[289]

The blameworthiness standard serves several functions in this dispute. First, some types of motivation in transgender cases can be clearly identified as blameworthy, such as transphobia or the enforcement of traditional gender roles. On the remaining issue of sexual deception, commentators have a genuine normative disagreement about whether a sexual deception of this type is a sufficient ground for responding in extreme anger. The blameworthiness standard does not itself resolve this dispute, and I do not purport to resolve it here. Nevertheless, the standard correctly frames the normative question: is a response of extreme anger, due to being deceived about the biological sex of one's intimate partner, an appropriate or a blameworthy emotional response?

CONCLUSION

The current provocation doctrine suffers from a deep conceptual problem—a failure to identify which features of a particular defendant are properly relevant when assessing the adequacy of provocation, and which features are not relevant. Commentators to date have failed to "provide a principled theory by which to distinguish those characteristics with which the law should individualize from those with which it should not."[290]

This article identifies the general solution to this puzzle: provocation is not reasonable if the reason the defendant became extremely angry is due to some blameworthy belief or attribute of the defendant. This principle—blameworthiness—distinguishes those aspects of the defendant that cannot form a basis for him to argue he was reasonably provoked from those aspects that can properly form the basis of a provocation claim.

---

provocation should not countenance the use of violence to enforce gender norms."); Steinberg, *supra* note 20, at 501–02.

286. Bigler, *supra* note 284, at 785–86.

287. *Id.*

288. Steinberg, *supra* note 20, at 509. Steinberg argues that there is no duty to reveal one's sex in a sexual encounter, and "[b]arring the existence of a duty to reveal one's sex, courts should not base mitigation instructions on an imagined breach of this contrived duty." *Id.* at 512.

289. Lee & Kwan, *supra* note 20, at 119.

290. Robinson, *supra* note 15, at 320.

The "blameworthiness" principle resolves, at a conceptual level, disputes over which cultural or other traits of a defendant should be "imported" onto the "reasonable person." It does so by reframing the issue: the jury should take the defendant as he is, considering all of his beliefs, traits, and attributes, and then determine whether the reason the defendant became extremely angry is blameworthy. This same principle identifies the correct normative standard against which to judge provocation claims in recurring, controversial contexts, such as men who kill their intimate partners, persons who kill in response to a same-sex sexual advance, or persons who kill after discovering their intimate partner is transgender. In each case, the decision-maker (whether judge, jury, or legislature) should focus on determining why the defendant became extremely angry and reject any reasons that are normatively blameworthy.

A defendant who pleads provocation asks the community to mitigate his wrongful act of killing from murder to manslaughter, and to do so because another person provoked him into a rage that made it much more difficult to control his violent response. But if the defendant's reason for becoming angry is blameworthy, he has no rightful claim to mitigation—he is culpably responsible for entering the state of extreme anger that might otherwise reduce his culpability.

The principle of "blameworthiness" does not provide ready solutions in all cases. In a diverse, pluralistic society, individuals will disagree about which attributes or reasons are blameworthy. Those disagreements might be resolved (or partly resolved) by legislatures or by courts, or left to juries. But the principle of blameworthiness solves the conceptual puzzle and provides the proper normative framework for channeling these disputes.