

# NOT FROM A WICKED HEART: TESTING THE ASSUMPTIONS OF THE PROVOCATION DOCTRINE

Carlton J. Patrick\* and Debra Lieberman\*\*

*The normative logic of the provocation doctrine rests on the long-held, yet untested, assumption that anger can motivate people to act in ways which they believe are morally wrong. Here we provide a frontline inquiry into this premise in the context of the quintessential provocation scenario: a man witnessing or learning of his partner's infidelity. Among men who had discovered a partner's affair, anger was more strongly correlated with motivation to retaliate than with judgments as to whether such retaliation was morally acceptable. Moreover, anger explained increases in motivation beyond what could be accounted for by increases in moral judgments. However, these effects were not uniform to all behaviors: anger motivated retaliation beyond what participants thought was morally acceptable only for those acts salient to the function of anger in this context (yelling, pushing, and striking). Taken together, these results partially support the traditional assumptions of the provocation doctrine while calling other aspects of the doctrine's normative framework into question.*

## TABLE OF CONTENTS

INTRODUCTION .....	34
I. THE PROVOCATION DOCTRINE .....	35
A. <i>Development and Variations in Brief</i> .....	35
B. <i>Normative Logic</i> .....	38
II. PRIOR RESEARCH: MORAL CONSCIENCE, MOTIVATION, AND EMOTIONS .....	43
III. THE PRESENT STUDY .....	48
A. <i>Methods</i> .....	48

---

\* Researcher with the Evolution and Human Behavior Laboratory, Ph.D. candidate at the University of Miami, M.S., University of Miami; J.D., Boston University School of Law; B.S., Florida State University. Address correspondence to carlton.patrick@gmail.com.

\*\* Professor of Psychology, University of Miami; Ph.D., University of California Santa Barbara; B.S. Binghamton University. The authors would like to thank Owen Jones, Oliver Goodenough, Mike McCullough, Tamara Lave, and the participants in the Society for Evolutionary Analysis in Law Scholarship Conference XVI in Tempe, Arizona for helpful comments on both earlier versions of this Article and presentations of the research presented herein. This research was made possible, in part, by a grant from the John Templeton Foundation.

1. <i>Participants</i> .....	48
2. <i>Materials and Procedure</i> .....	48
B. <i>Analysis and Results</i> .....	50
IV. DISCUSSION .....	54
A. <i>Study Limitations</i> .....	54
B. <i>Normative Implications</i> .....	56
CONCLUSION .....	58

#### INTRODUCTION

In most jurisdictions in the United States, homicides committed in the “heat of passion” are treated as manslaughter and are punished less severely than premeditated or intentional killings.<sup>1</sup> In some jurisdictions, the punishment for assault or battery can be similarly mitigated.<sup>2</sup> The heat of passion doctrine—sometimes referred to as the doctrine of provocation<sup>3</sup>—is premised on the assumption that one who kills while in a highly emotional state may not act from “any wickedness of heart or cruelty or recklessness of disposition,” but rather as “[a] result of the temporary excitement, by which the control of reason was disturbed.”<sup>4</sup>

Importantly, this rationale assumes a certain feature of human psychology: that being in a highly emotional state can cause people to behave in ways that they otherwise feel are morally unacceptable.<sup>5</sup> This phenomenon has, surprisingly, received little attention from behavioral researchers. There has been much work focusing on the relationship between emotion states and behavior,<sup>6</sup> and separately on the influence of emotions on moral judgment,<sup>7</sup> but there has been very little research that has attempted to triangulate the relationship between the three. As a result, the provocation doctrine—a doctrine that can be traced as far back as twelfth-century English decisions—remains reasoned on a

<sup>1</sup> Mitchell N. Berman & Ian P. Farrell, *Provocation Manslaughter as Partial Justification and Partial Excuse*, 52 WM. & MARY L. REV. 1027, 1031 n.1 (2011) (explaining that the provocation defense exists in one form or another in almost every state, and noting that in Illinois and Texas provocation results in a lower category of murder rather than murder).

<sup>2</sup> See, e.g., *Christian v. State*, 951 A.2d 832, 848 (Md. 2008) (holding “that the mitigation defenses of hot-blooded response to adequate provocation and imperfect self-defense could apply to [cases of] first degree assault . . .”); *State v. Smith*, 858 N.E.2d 1222, 1235 (Ohio Ct. App. 2006) (discussing the mitigating circumstance of provocation for cases of aggravated assault).

<sup>3</sup> We use the terms interchangeably here.

<sup>4</sup> E.g., *People v. Beltran*, 301 P.3d 1120, 1128 (Cal. 2013) (citing *Maier v. People*, 10 Mich. 212, 219 (Mich. 1862)).

<sup>5</sup> *Infra* notes 28–30 and accompanying text.

<sup>6</sup> See generally Jennifer S. Lerner et al., *Emotion and Decision Making*, 66 ANN. REV. PSYCHOL. 799 (2015) (reviewing psychological research on emotions and behavior).

<sup>7</sup> See generally Yana R. Avramova & Yoel Inbar, *Emotion and Moral Judgment*, 4 WILEY INTERDISC. REVIEWS: COGNITIVE SCI. 169 (2013) (reviewing psychological research on emotions and moral judgment).

largely-untested psychological premise.<sup>8</sup> This Article provides a frontline examination of this premise by testing whether anger—the emotion most associated with the provocation doctrine—can cause people to act in ways that they feel are otherwise morally unacceptable.

## I. THE PROVOCATION DOCTRINE

### A. *Development and Variations in Brief*

While an exhaustive overview of the historical development and various iterations of the provocation doctrine is outside the scope of this article, a brief review should help place the assumptions being tested here in context.<sup>9</sup>

The United States adopted the provocation doctrine from English common law, but variants of the rule can be found in other, much older, legal codes.<sup>10</sup> The Attic Code of Ancient Greece, for example, specified that the usual punishment of exile could be mitigated for certain killings including “kill[ing] another unintentionally in an athletic contest, or overcoming him in a fight on the highway, or unwittingly in battle, or in intercourse with his wife, or mother, or sister, or daughter, or concubine kept for the procreation of legitimate children.”<sup>11</sup> The first recognizable formation of the provocation standard, which appears to have emerged in the seventeenth century in England, required the killing to have occurred while the killer was in the “‘heat of blood’” and that the heat of blood must have been brought about by provocation that was “sufficiently grave.”<sup>12</sup> And by the late nineteenth century, the United States Supreme Court had announced the rule that, to a large degree, resembles the framework in place today:

---

<sup>8</sup> See JEREMY HORDER, PROVOCATION AND RESPONSIBILITY 1–22 (1992) (discussing the historical development of the provocation doctrine).

<sup>9</sup> Over the past thirty-five years, legal scholars have produced a spate of articles detailing the doctrinal evolution of provocation laws. By and large, these accounts have had few points of disagreement; recent articles tracing the historical development of the doctrine tend to coalesce on the same general account, and moreover tend to cite many of the same primary sources for information. For some in-depth accounts see, e.g., *id.* at 1–42; Berman & Farrell, *supra* note 1, at 1035–44; Joshua Dressler, *Rethinking Heat of Passion: A Defense in Search of a Rationale*, 73 J. CRIM. L. & CRIMINOLOGY 421, 425–32 (1982) [hereinafter Dressler, *Rethinking Heat of Passion*]. See generally SAMUEL H. PILLSBURY, JUDGING EVIL: RETHINKING THE LAW OF MURDER AND MANSLAUGHTER (1998); A.J. Ashworth, *The Doctrine of Provocation*, 35 CAMBRIDGE L.J. 292 (1976).

<sup>10</sup> Theophile J. Meek (trans.), *The Middle Assyrian Laws, in ANCIENT NEAR EASTERN TEXTS RELATING TO THE OLD TESTAMENT* 181 (James B. Pritchard ed., 3d ed., 1969).

<sup>11</sup> DAVID COHEN, LAW, SEXUALITY, AND SOCIETY: THE ENFORCEMENT OF MORALS IN CLASSICAL ATHENS 100 (1991) (“If a man kills another unintentionally in an athletic contest, or overcoming him in a fight on the highway, or unwittingly in battle, or in intercourse with his wife, or mother, or sister, or daughter, or concubine kept for procreation of legitimate children, he shall not go into exile as a manslayer on that account.”) (quoting Demosthenes’ oration *Against Aristocrates*).

<sup>12</sup> Berman & Farrell, *supra* note 1, at 1035–36 (outlining the crystallization of the contemporary standard in 17th century England) (citations omitted).

The law in recognition of the frailty of human nature, regards a homicide committed under the influence of sudden passion, or in hot blood, produced by adequate cause, and before a reasonable time has elapsed for the blood to cool, as an offense of a less heinous character than murder.<sup>13</sup>

Historically, the law defined specific, narrow categories as to what could count as adequate provocation.<sup>14</sup> These traditional categories had their roots in the male code of honor, and were thought to be such an affront to one's reputation as to compel a response.<sup>15</sup> They are often referred to as the "nineteenth century four," and include: (1) a "grossly insulting assault," (2) "witnessing an attack upon a friend or relative," (3) witnessing a false arrest, and (4) "witnessing one's wife in the act of adultery."<sup>16</sup> Adultery, specifically, has always occupied a privileged status at the heart of the doctrine.<sup>17</sup> In fact, nearly all scholars agree that the paradigmatic case of provocation is that of a man who, upon discovering his spouse *in flagrante delicto*, flies into a violent rage and kills his partner, her paramour, or both.<sup>18</sup>

<sup>13</sup> *Andersen v. United States*, 170 U.S. 481, 510 (1898).

<sup>14</sup> Stephen P. Garvey, *Passion's Puzzle*, 90 IOWA L. REV. 1677, 1688–89 (2005) ("At common law, for example, an alleged provocation was inadequate as a matter of law if it fell outside certain narrow categories of adequacy, and was adequate as a matter of law if it fell within them.").

<sup>15</sup> See HORDER, *supra* note 8, at 25–29 (explaining the doctrine as a product of honor theory); Victoria Nourse, *Passion's Progress: Modern Law Reform and the Provocation Defense*, 106 YALE L.J. 1331, 1340–41 (1997) (tracing the traditional categories to a definition of provocation "derived from an older social order, indeed, a code of honor") (citations omitted); see also Donna K. Coker, *Heat of Passion and Wife Killing: Men Who Batter/Men Who Kill*, 2 S. CAL. REV. L. & WOMEN'S STUD. 71, 79 (1992) (tracing the doctrine to the settlement of "breaches of honor").

<sup>16</sup> See Berman & Farrell, *supra* note 1, at 1036 (outlining the "four distinct—and exhaustive—categories of provocative conduct considered 'sufficiently grave to warrant the reduction from murder to manslaughter of a hot-blooded intentional killing.'"); see also Coker, *supra* note 15, at 80 (citing M.D.G., Note, *Manslaughter and the Adequacy of Provocation: The Reasonableness of the Reasonable Man*, 106 U. PA. L. REV. 1021, 1023–24 (1958), on "nineteenth century four").

<sup>17</sup> *Compare* *People v. Beltran*, 301 P.3d 1120, 1128 (Cal. 2013) (citing *Manning's Case*, 83 ENG. REP. 112 (1670)) (where the judge sentenced the defendant—who killed his wife after catching her in the throes of an affair—to a burning of the hand but directing the executioner to "burn him gently, because there could not be greater provocation than this"), with Tamar Lewin, *What Penalty for a Killing in Passion?*, N.Y. TIMES (Oct. 21, 1994), <http://www.nytimes.com/1994/10/21/us/what-penalty-for-a-killing-in-passion.html?mcubz=1> [<https://perma.cc/EKX6-6H8S>] (where the judge sentenced the defendant—who killed his wife after finding her in bed with another man—to 18 months in prison saying he wished he did not have to send him to prison at all, and wondering "how many men married five, four years would have the strength to walk away without inflicting some corporal punishment"). As a frame of reference, in the span of time between these pronouncements, the world witnessed the discoveries of gravity, electricity, evolution, germ theory, x-rays, penicillin and DNA, not to mention Sigmund Freud, William James, the Cognitive Revolution, neuroscience and nearly every modern theory of human behavior. See generally JOHN HENRY, A SHORT HISTORY OF SCIENTIFIC THOUGHT (2012).

<sup>18</sup> See, e.g., Coker, *supra* note 15, at 72 ("English and American jurists and legal scholars repeatedly refer to adultery as the paradigm example of provocation adequate enough to mit-

Over time, these strict nineteenth century categories have been replaced by broader standards that encapsulate a range of potential provoking events.<sup>19</sup> The typical common law standard, for example, is structured such that any number of circumstances might constitute adequate provocation so long as the defendant can show: (1) that a reasonable person in the defendant's situation would have been adequately provoked, (2) that the defendant was in-fact provoked and became emotionally charged to an extent that he lost self-control, (3) that a reasonable person in the defendant's situation would not have had sufficient time to "cool off" between provocation and killing, and (4) that the defendant did not, in-fact, cool off before killing his victim.<sup>20</sup> In jurisdictions that have adopted the heat of passion standards recommended by the Model Penal Code, the umbrella of killings covered by the defense has been broadened even further to include homicides "committed under the influence of extreme mental or emotional disturbance for which there is a reasonable explanation or excuse."<sup>21</sup>

In addition to the expansion of what might constitute "adequate provocation," modern doctrines have also expanded what it means to be "provoked" or "in the heat of passion." Historically, passion was treated as synonymous with

---

igate what would otherwise be murder to a voluntary manslaughter conviction."); Antonia Elise Miller, Note, *Inherent (Gender) Unreasonableness of the Concept of Reasonableness in the Context of Manslaughter Committed in the Heat of Passion*, 17 WM. & MARY J. WOMEN & L. 249, 257 (2010) ("In today's society . . . the typical paradigm for the use of voluntary manslaughter is exemplified by the facts of the *Peacock* case: an unsuspecting husband comes home and catches his wife in bed with another man."). See also Susan D. Rozelle, *Controlling Passion: Adultery and the Provocation Defense*, 37 RUTGERS L.J. 197, 198–99 (2005) ("These 'nineteenth century four' reduce to two paradigmatic contexts, still the paradigms today: (1) the defendant, faced with a violent but non-deadly assault, responds to his attacker in the heat of passion with excessive, deadly force; and (2) the defendant, arriving home early one day to discover his wife in bed with another man, in the heat of passion kills one, the other, or both.") (citations omitted).

<sup>19</sup> Aya Gruber, *A Provocative Defense*, 103 CAL. L. REV. 273, 280 (2015) (discussing the expansion of the provocation standard) (citations omitted).

<sup>20</sup> See WAYNE R. LAFAVE, *CRIMINAL LAW* 777 (4th ed. 2003) (outlining the four requirements); Reid Griffith Fontaine, *Adequate (Non)Provocation and Heat of Passion as Excuse Not Justification*, 43 U. MICH. J.L. REFORM 27, 29–30 (2009); see also Berman & Farrell, *supra* note 1, at 1040–42 (arguing that these four elements are just other ways of stating the only two essential requirements: that "[t]he actor must have killed while in the *heat of passion*" and that "[t]he heat of passion must have been brought about by *adequate provocation*").

<sup>21</sup> MODEL PENAL CODE § 210.3(1)(b) (1980); Joshua Dressler, *Why Keep the Provocation Defense?: Some Reflections on a Difficult Subject*, 86 MINN. L. REV. 959, 960 (2002) [hereinafter Dressler, *Why Keep the Provocation Defense?*] ("In 1962, the drafters of the Model Penal Code provided a new, far broader vision of the defense."); Berman & Farrell, *supra* note 1, at 1044 ("The Model Penal Code's EMED [extreme mental or emotional disturbance] formulation represents a substantial reform of the provocation defense. The EMED defense is not restricted to loss of self-control caused by passion stemming from adequate provocation. The disturbance undermining self-control may be mental as well as emotional, and the emotional or mental disturbance need not arise from provoking conduct at all. When provocation occurs, the EMED defense may apply regardless of whether the person killed was the provoker, or whether the provocation was directed at the defendant or a third party.").

intense anger or rage.<sup>22</sup> Today, however, in common law jurisdictions, passion is often broad enough to include emotions like fear or resentment.<sup>23</sup> In jurisdictions that follow the Model Penal Code, passion can include any “mental or emotional disturbance,” so long as it is “extreme.”<sup>24</sup> The operative limitation is often now one of objectivity: would the emotion in question cause an “‘ordinary person of average disposition’ ” to act “‘from passion rather than judgment’ ”?<sup>25</sup> Still, despite the general expansion of the definition of passion, most scholars agree that anger remains the quintessential case.<sup>26</sup>

### B. Normative Logic

The normative judgment embedded in the various iterations of the provocation doctrine is that an individual motivated by intense emotion into committing an act they understand to be wrong is, while not blameless, less deserving of punishment than someone who acts with cool, deliberative intent.<sup>27</sup> Courts sometimes characterize the difference as acting “due to anger, not evilness.”<sup>28</sup> The comments to the Model Penal Code frame it this way:

---

<sup>22</sup> See, e.g., *People v. Sica*, 245 P. 461, 463 (Cal. Dist. Ct. App. 1926) (“‘Anger’ and ‘passion’ are interchangeable, and mean practically the same thing.”) (citation omitted); Nourse, *supra* note 15, at 1341 (noting that the original four categories of the nineteenth century were “thought to isolate those cases in which the violation of a social or relational norm led to righteous anger.”).

<sup>23</sup> See, e.g., *People v. Wu*, No. E007993, 286 Cal. Rptr. 868, 884 (Cal. Ct. App. 1991) (noting that “‘passion’ need not mean ‘rage’ or ‘anger’ but may be any ‘[v]iolent, intense, high-wrought or enthusiastic emotion’”) (quoting *People v. Borchers*, 325 P.2d 97, 102 (Cal. 1958)).

<sup>24</sup> See Coker, *supra* note 15, at 79 (“‘Passion’ has usually meant ‘anger,’ though some jurisdictions have consistently included fear and jurisdictions that follow the Model Penal Code include any ‘extreme mental or emotional disturbance.’”).

<sup>25</sup> Coker, *supra* note 15, at 79 (citing 17 Cal. Jur. 3d (Rev.) Part 1, Criminal Law § 255 (1984)).

<sup>26</sup> See Coker, *supra* note 15, at 102 (“If adultery is the ‘paradigm’ heat of passion event, anger is the paradigm heat of passion emotion.”); Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 427 n.62 (“[A]nger is the usual emotion alleged in provocation cases . . .”); Dressler, *Why Keep the Provocation Defense?*, *supra* note 21, at 959 n.5 (“Provocation law is all about emotions, most notably anger.”); Garvey, *supra* note 14, at 1687 n.34 (“The passion or emotion . . . implicated in cases of provocation is usually thought to be anger or resentment . . .”); Laurie J. Taylor, Comment, *Provoked Reason in Men and Women: Heat-of-Passion Manslaughter and Imperfect Self-Defense*, 33 UCLA L. REV. 1679, 1711–12 (1986) (“[R]age is still the paradigm emotion for heat of passion . . .”).

<sup>27</sup> See Nikolette Y. Clavel, *Righting the Wrong and Seeing Red: Heat of Passion, the Model Penal Code, and Domestic Violence*, 46 NEW ENG. L. REV. 329, 333 (2012) (“The doctrine of common law voluntary manslaughter evolved from the recognition that those who kill in the heat of passion are less culpable than those who commit premeditated murders.”) (citations omitted); Fontaine, *supra* note 20, at 49 (“It is not the provocation that mitigates the defendant’s culpability and punishment, but the emotionally-charged effect that is has on him.”).

<sup>28</sup> Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 462.

[The provocation doctrine] is a concession to human weakness and perhaps to non-deterability, a recognition of the fact that one who kills in response to certain provoking events should be regarded as demonstrating a significantly different character deficiency than one who kills in their absence . . . The underlying judgment is thus that some instances of intentional homicide may be as much attributable to the extraordinary nature of the situation as to the moral depravity of the actor.<sup>29</sup>

This normative judgment, in turn, relies on an implicit assumption of human psychology: that intense emotion can cause an individual to deviate from what they believe is morally acceptable—to cause them to do something that they understand to be wrong and, presumably, that they would refrain from doing *but for* the emotion.<sup>30</sup>

While there may be a general scholarly consensus as to descriptive development of the doctrine, there is extensive normative debate as to exactly why (let alone, whether), this “frailty of human nature” should mitigate the actor’s punishment. Entire symposiums, for example, have been staged around the question of whether intense emotion partially *justifies* the actor’s behavior (i.e., entitles the actor to a certain degree of retaliation), or partially *excuses* it (i.e., renders the actor less morally blameworthy).<sup>31</sup> Other debates have centered on whether the mitigation is best understood under a retributive framework (e.g., the passionate actor is less deserving of punishment than the deliberative actor) or a utilitarian one (e.g., the passionate actor is less deterrable than the deliberative actor).<sup>32</sup> Even within the retributive framework—the favored explanation of the doctrine—there are different rationales as to why the actor is less deserv-

---

<sup>29</sup> MODEL PENAL CODE § 210.3 cmt. 5(a) (1980).

<sup>30</sup> See, e.g., Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 464 (“In short, provocation is an excuse premised upon involuntariness based upon reduced choice-capabilities.”); Fontaine, *supra* note 20, at 48 (“Heat of passion reflects a mediated path by which there is an interpretation of provocation (cognition) that causes overwhelming anger (emotion), which, in turn, leads to a reactive killing (behavior).”); Garvey, *supra* note 14, at 1713 (analogizing crimes of passion to hate crimes and explaining that “[the actor’s] desire and belief are but-for causes of his action, but they are not, so it might be put, *motives* for his action.”).

<sup>31</sup> See generally Symposium, *The Nature, Structure, and Function of Heat of Passion/Provocation as a Criminal Defense*, 43 U. MICH. J.L. REFORM 1 (2009); Vera Bergelson, *Justification or Excuse? Exploring the Meaning of Provocation*, 42 TEX. TECH L. REV. 307 (2009).

<sup>32</sup> Compare Dressler, *Why Keep the Provocation Defense?*, *supra* note 21, at 963 (“[M]ost modern scholars would agree that the basis for the defense of provocation is found in retributive concepts of desert . . .”) with Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 464 (“Anger makes us less *able* to respond in a legally and morally appropriate fashion.”). See also MODEL PENAL CODE § 210.3 cmt. 5(a) (1980), *supra* note 29 and accompanying text (which combines both rationales: “[The provocation doctrine] is a concession to human weakness and perhaps to non-deterability . . .”); Marcia Baron, *The Provocation Defense and the Nature of Justification*, 43 U. MICH. J.L. REFORM 117, 127 (2009) (combining elements of both in noting that “[i]t is too harsh [to blame the individual] either because of some feature of the agent for which she is not culpable that rendered it very difficult for her to act as she should, or because the circumstances were such that it would have been very difficult for most people to act as they should.”) (citation omitted).

ing of punishment. Is it because he was incapable of forming the blameworthy moral judgment—the specific intent to kill, or malice aforethought—usually required for murder?<sup>33</sup> Is it because he was, for the duration of the passion, dethroned of his reasoning faculties, and was thus acting “not out of rational thought but out of unconsidered reaction to the provocation”?<sup>34</sup> Or is it because the actor has “demonstrated fewer blameworthy character flaws than are usually observed in killers?”<sup>35</sup> In short, the doctrine remains, as Dressler remarked more than three decades ago, “a [defense] in search of a . . . rationale.”<sup>36</sup>

These arguments, however, have simply bypassed the threshold question of whether intense emotion—typically, anger—can actually motivate an individual to behave in ways that they otherwise believe to be morally wrong. While it may seem common sense that anger motivates us to do things we understand to be wrong,<sup>37</sup> discoveries in both the natural and behavioral sciences have repeatedly shown that our common sense experience is often misleading and inaccurate: we observe the world as flat and motionless instead of round and revolving; we perceive motion in absolute terms instead of relative terms; we give good-faith explanations for our behavior that are demonstrably inaccurate; we “remember” events that never happened; and we consider ourselves rational decision-makers in spite of the fact that we frequently make sub-optimal choices.<sup>38</sup> In other words, we cannot just rely on our intuitions if we want to truly understand the psychological mechanics that underlie our experience of being “in the heat of passion.” Empirical verification is necessary.

Importantly, an empirical examination of the doctrine holds value regardless of whether the assumptions embedded in the doctrine are accurate. On the one hand, if a state of intense anger does not in fact produce the assumed effect—that is, motivate the actor to act in a way that he or she believes morally unacceptable—then many of the normative debates that follow the doctrine are proceeding largely in vain. On the other hand, even if anger does produce the

---

<sup>33</sup> See, e.g., *State v. Lee*, 321 N.W.2d 108, 110 (Wis. 1982) (citing Wisconsin Criminal Jury Instruction #1130, which states that heat of passion must make the actor “incapable of forming and executing that distinct intent to take human life essential to murder in the first degree”).

<sup>34</sup> *People v. Beltran*, 301 P.3d 1120, 1125 (Cal. 2013).

<sup>35</sup> Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 462.

<sup>36</sup> *Id.* at 432 (calling heat of passion “a doctrine in search of a modern-day rationale.”).

<sup>37</sup> See, e.g., Dressler, *Rethinking Heat of Passion*, *supra* note 9, at 463, 463–64 n.299 and accompanying text (declaring that “I do not rely on science in this article to inform the analysis regarding the nature of anger as it affects human conduct” and noting instead that he relies on “our common experience.”); Stephen J. Morse, *The Irreducibly Normative Nature of Provocation/Passion*, 43 U. MICH. J.L. REFORM 193, 194 (2009) (“Most rational adults understand two aspects of our psychological natures. First, things that happen can make people extremely angry. Second, we have much more difficulty behaving rationally when we are in states of such extreme emotion or other states that diminish rationality.”).

<sup>38</sup> See Carlton J. Patrick, *The Long-Term Promise of Evolutionary Psychology for the Law*, 48 ARIZ. ST. L.J. 995, 1007–12 (2016) (discussing and giving examples of this type of “instinct blindness”).



assumed effect, a host of tributary questions remain: is the effect uniform for all behaviors, and all moral judgments, or is the actor able to act from “judgment” and not “passion” in some contexts? Does anger only produce a change in behavioral motivations, or does it also affect moral judgments as well? Asked another way: does the heat of passion cause the individual not just to behave in a way in which he ordinarily would not, but also to believe his actions are more justifiable? Is the effect limited to instances of anger or rage, or does a more modern conception, such as the Model Penal Code’s allowance of any extreme mental or emotional disturbance, better capture the underlying psychological impetus?

Unfortunately for lawmakers and legal scholars, the existing empirical record is inconclusive. Despite decades of psychological research on anger, we are unaware of any study that directly tests the effect supposed by the doctrine. Existing theory and evidence do suggest that anger functions to orchestrate aggression and punishment in response to the undervaluing of one’s welfare by another person.<sup>39</sup> In other words, anger serves a recalibrational function—to cause the individual towards whom the anger is directed to increase the value that they place on the angry individual’s welfare.<sup>40</sup> In line with this recalibrational model, a large body of research demonstrates a general escalatory effect of anger on motivations to retaliate. Research shows, for example, that anger increases both the likelihood<sup>41</sup> and severity<sup>42</sup> of punishment, that it enhances

---

<sup>39</sup> See, e.g., NICO H. FRIJDA, *THE EMOTIONS* 198 (1986) (describing the arousal of anger when a person has the experience of being slighted or hurt by intentional acts of another person); Aaron Sell et al., *Formidability and the Logic of Human Anger*, 106 *PROC. NAT’L ACAD. SCI.* 15073, 15074 (2009) (“[W]hen the anger program detects that the other party is not placing ‘sufficient’ weight on the welfare of the actor, anger is triggered.”).

<sup>40</sup> See, e.g., Sell et al., *supra* note 39, at 15073 (“[T]he function of anger is to orchestrate behavior in the angry individual that creates incentives in the target of the anger to recalibrate upwards the weight he or she puts on the welfare of the angry individual.”); Agneta H. Fischer & Ira J. Roseman, *Beat Them or Ban Them: The Characteristics and Social Functions of Anger and Contempt*, 93 *J. PERSONALITY & SOC. PSYCHOL.* 103, 104 (2007) (“[T]he social function of anger can be conceptualized as attaining a better outcome by forcing a change in another person’s behavior. This function can be served by hostile or antagonistic behaviors, that is, by seeking confrontation or by attacking someone”).

<sup>41</sup> See, e.g., Julie H. Goldberg et al., *Rage and Reason: The Psychology of the Intuitive Prosecutor*, 29 *EUR. J. SOC. PSYCHOL.* 781, 781 (1999) (finding that anger aroused by the belief that an injustice occurred “lower[ed] . . . thresholds for making attributions of harmful intent”); Madan M. Pillutla & J. Keith Murnighan, *Unfairness, Anger, and Spite: Emotional Rejections of Ultimatum Offers*, 68 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 208, 215–19 (1996) (finding that anger predicted retaliation in ultimatum game experiments).

<sup>42</sup> See, e.g., Leah C. Georges et al., *The Angry Juror: Sentencing Decisions in First-Degree Murder*, 27 *APPLIED COGNITIVE PSYCHOL.* 156, 160–63 (2013) (finding that as mock jurors’ anger increased, the more likely they were to assign a death sentence, and the weaker their ratings of mitigating factors).

motivation for social dominance,<sup>43</sup> increases a tolerance for risk,<sup>44</sup> and increases a willingness to endure additional personal costs in order to mete out punishment.<sup>45</sup>

The question of how this well-established effect interacts with the angry individual's moral judgments, however, is a novel one. The overarching shortcoming of the existing literature is that studies generally include only two of the three variables of interest (i.e. anger, motivation, and moral judgment). Studies examining the motivational effects of anger, for example, have typically not included measures of how morally acceptable (or unacceptable) participants believe their actions to be.<sup>46</sup> Studies focusing on the effects of anger on moral judgment have likewise not included independent measures of motivation or behavior,<sup>47</sup> and moreover tend to focus on how anger influences moral judgments as to the actions of others,<sup>48</sup> or of behavior or punishment in the abstract.<sup>49</sup> Because moral judgments as to the behavior of others may be generated by a set of psychological mechanisms that are distinct from those used to generate moral judgments as to the behavior of the self, however, it can be problematic to generalize from one class of studies to the other.<sup>50</sup>

---

<sup>43</sup> See generally Larissa Z. Tiedens, *Anger and Advancement Versus Sadness and Subjugation: The Effect of Negative Emotion Expressions on Social Status Conferral*, 80 J. PERSONALITY & SOC. PSYCHOL. 86 (2001).

<sup>44</sup> See Daniel M.T. Fessler et al., *Angry Men and Disgusted Women: An Evolutionary Approach to the Influence of Emotions on Risk Taking*, 95 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 107, 116 (2004) (finding that anger increases risk taking in men); Jennifer S. Lerner & Dacher Keltner, *Fear, Anger, and Risk*, 81 J. PERSONALITY & SOC. PSYCHOL. 146, 146 (2001) (explaining the finding that "angry people expressed optimistic risk estimates and risk-seeking choices.").

<sup>45</sup> See generally Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137 (2002); Elise C. Seip et al., *Anger Motivates Costly Punishment of Unfair Behavior*, 38 MOTIVATION & EMOTION 578 (2014).

<sup>46</sup> See, e.g., *supra* notes 39–45.

<sup>47</sup> See, e.g., Evan Polman & Rachel L. Ruttan, *Effects of Anger, Guilt, and Envy on Moral Hypocrisy*, 38 PERSONALITY & SOC. PSYCHOL. BULL. 129, 133 (2012) (angered participants rating the acceptability of hypothetical moral transgressions instead of measuring their willingness or motivations to transgress).

<sup>48</sup> See, e.g., Matúš Grežo & L'ubor Pilárik, *Anger and Moral Reasoning in Decision Making*, 4 J. EUR. PSYCHOL. STUDENTS 56, 62 (2013) (evaluating participants' anger and moral judgments with respect to punishment options in the abstract, not with respect to how participants own punishment decisions).

<sup>49</sup> See, e.g., Jessica M. Salerno & Liana C. Peter-Hagene, *The Interactive Effect of Anger and Disgust on Moral Outrage and Judgments*, 24 PSYCHOL. SCI. 2069, 2070–71 (2013) (discussing participants' assessment of anger and moral outrage toward defendant in vignettes).

<sup>50</sup> See Kelly Asao & David M. Buss, *The Tripartite Theory of Machiavellian Morality: Judgment, Influence, and Conscience as Distinct Moral Adaptations*, in THE EVOLUTION OF MORALITY 6–7 (Todd K. Shackelford & Ranald D. Hansen eds., 2016) (outlining the difference between the psychological mechanisms and moral judgments generated with respect to the self on one hand and conspecifics on the other); Peter DeScioli & Robert Kurzban, *Mysteries of Morality*, 112 COGNITION 281, 285–86 (2009) (explaining the separate psychologi-

A few studies have attempted to examine how anger and related emotions influence decisions to engage in “unethical” or “immoral” behavior, but because these studies generally do not measure how morally acceptable the participants believe their own actions to be, they are also generally inconclusive.<sup>51</sup> As a result, we can only assume that the participants viewed the acts as immoral; we are unable to rule out the hypothesis that the angry individuals believed their behavior—as a result of their anger or otherwise—to be morally justified. In sum, existing research does not allow us to sufficiently answer the true question at the heart of the provocation doctrine: do angry individuals behave in ways that they otherwise believe to be wrong?

## II. PRIOR RESEARCH: MORAL CONSCIENCE, MOTIVATION, AND EMOTIONS

While there are no studies *directly* addressing the question at the heart of the doctrine, other existing research does allow us to make certain predictions regarding the three-way interaction between moral judgments, behavioral motivations, and emotions.

One prominent theory, for example, suggests that the function of our moral conscience—the subjective feeling that personally taking a particular action is, or would be, “right” or “wrong”—is to predict the probable judgment of others and guide behavior in order to avoid punishment and social condemnation.<sup>52</sup> In this view, our moral conscience can be thought of as a forecasting instrument that computes the expected social consequences of a particular action and then consciously delivers a moral judgment or intuition with regard to that action.<sup>53</sup> Expected positive outcomes (social praise, positive gossip, increases in reputation, elevated social standing) are encouraged by tagging prospective actions with sensations such as *right*, *good*, *virtuous*, or *justified*, while expected negative outcomes (revenge, negative gossip, loss of reputation, social ostracism,

---

cal functions of moral conscience and moral condemnation based on whether the judgment refers to behavior of the self or the behavior of others).

<sup>51</sup> See generally Rimma Teper et al., *How Emotions Shape Moral Behavior: Some Answers (and Questions) for the Field of Moral Psychology*, 9 SOC. & PERSONALITY PSYCHOL. COMPASS 1 (2015); cf. Dan Ariely & George Loewenstein, *The Heat of the Moment: The Effect of Sexual Arousal on Sexual Decision Making*, 19 J. BEHAV. DECISION MAKING 87, 94 (2006) (finding that sexual arousal increased participants’ willingness to engage in morally questionable behavior where “morally questionable” was defined by the authors—participants did not rate the morality of their actions).

<sup>52</sup> See Peter DeScioli & Robert Kurzban, *A Solution to the Mysteries of Morality*, 139 PSYCHOL. BULL. 477, 487–88 (explaining moral conscience as a mechanism for anticipation and avoiding condemnation by third parties); Asao & Buss, *supra* note 50, at 5 (“[M]oral conscience, is a set of psychological mechanisms designed to guide one’s own behavior toward others to avoid negative fitness consequences as a result of judgment and influence mechanisms in others.”); Dan Sperber & Nicolas Baumard, *Moral Reputation: An Evolutionary and Cognitive Perspective*, 27 MIND & LANGUAGE 495, 495 (2012) (explaining that, from an evolutionary point of view, the function of moral behavior may be to secure a good reputation as a co-operator).

<sup>53</sup> Sperber & Baumard, *supra* note 52, at 495.

and other forms of individual or community punishment) are discouraged by tagging prospective actions with sensations such as *wrong*, *bad*, *shameful*, or *unjustified*.<sup>54</sup> Accordingly, humans should, all other things being equal, generally behave in accordance with their moral conscience—that is, people should typically be motivated to do what they feel is right and to refrain from what they feel is wrong.<sup>55</sup>

Often, however, all other things are not equal. In some situations, the overall advantages<sup>56</sup> of engaging in a particular behavior might outweigh the expected negative social consequences: when there is little chance of a moral transgression being detected,<sup>57</sup> when the individual is in a position of power and can “afford” reputational harm,<sup>58</sup> when the individual is already a social outcast and has little to fear from social devaluation,<sup>59</sup> when complying means

---

<sup>54</sup> See DeScioli & Kurzban, *supra* note 50, at 287 (“Conscience, on the present view, is a cognitive system that uses the moral concepts ‘right’ and ‘wrong’ to guide actor behavior.”); Sperber & Baumard, *supra* note 52, at 509 (Discussing the concern for one’s reputation as an intuitive mechanism guiding “the moral sense,” which “provides intuitions about what is right and wrong and involves moral sentiments with motivating power.”).

<sup>55</sup> See, e.g., Sperber & Baumard, *supra* note 52, at 497 (“Moral emotions are a signal of trustworthiness only to the extent that the underlying moral judgment is itself trustworthy and actually guides the individual’s behaviour.”). See generally Jerome Kagan, *Human Morality and Temperament*, in NEBRASKA SYMPOSIUM ON MOTIVATION: VOL. 51. MORAL MOTIVATION THROUGH THE LIFE SPAN (Gustavo Carlo & Carolyn Pope Edwards eds., 2005).

<sup>56</sup> We use the terms “advantages,” “benefits,” and “payoffs” in their evolved psychological sense—meaning, advantages/benefits/payoffs in terms of increased survival and reproduction over evolutionary time. We do not use these terms in the way they are commonly used and understood in an economic sense, to convey net payoffs measured from a contemporary perspective of economic rationality. In many cases, we should expect the two to differ. See Carlton J. Patrick, *A New Synthesis for Law and Emotions: Insights from the Behavioral Sciences*, 47 ARIZ. ST. L.J. 1239, 1271–77 (2015) for a discussion of this distinction in terms of emotions and the law. See Leda Cosmides & John Tooby, *Better than Rational: Evolutionary Psychology and the Invisible Hand*, 84 AM. ECON. REV. 327 (1994) for an explanation of this idea, generally.

<sup>57</sup> See, e.g., DeScioli & Kurzban, *supra* note 50, at 290 (covering evidence from studies that indicates moral conscience allows immoral behavior when condemnation is unlikely).

<sup>58</sup> See, e.g., Craig Haney et al., *A Study of Prisoners and Guards in a Simulated Prison*, 30 NAVAL RESEARCH REV. 1 (1973) (discussing a simulated prison experiment that divided participants randomly into prisoner and guard roles, and where the power granted to the guards resulted in aggression and cruelty towards the “prisoner” participants); Sell et al., *supra* note 39, at 15075–77 (finding that individuals with characteristics, such as strength in men or physical attractiveness in women, that render them less vulnerable to social devaluation are more prone to entitlement and to endorse the utility of personal aggression).

<sup>59</sup> See, e.g., MARTIN DALY & MARGO WILSON, *HOMICIDE* 124 (1988) (outlining the most common profile of killers in a study of Detroit homicides as “nobodies: unpropertied and unmarried, little educated, often unemployed.”); STEVEN PINKER, *THE BETTER ANGELS OF OUR NATURE: WHY VIOLENCE HAS DECLINED* 396 (2011) (“Among humans, the male may use coercion to get sex when certain risk factors line up: when he is violent, callous, and reckless by temperament; when he is a loser who cannot attract sexual partners by other means; when he is an outcast and has little fear of opprobrium from the community; and when he senses that the risks of punishment are low, such as during conquests and pogroms.”).

contravening authority, risking physical danger, or incurring other prohibitively high personal costs,<sup>60</sup> when other reputational losses, such as the loss of perception of formidability, outweigh the risk of condemnation,<sup>61</sup> or when the overall benefits of transgressing are so high that the net calculus favors risking any negative social consequences.<sup>62</sup> Thus, while the moral conscience may serve as one input for behavior, there are also many instances in which moral conscience and behavior might not match one-to-one.<sup>63</sup> In such instances, an individual might feel that a particular course of action would be the “right thing to do” and yet not be motivated to act, or conversely experience a willingness to engage in a behavior that the individual simultaneously believes to be “wrong.”<sup>64</sup>

Applying the view that emotions function to coordinate cognitive, behavioral, and physiological mechanisms to respond to certain recurrent classes of events, it follows that emotions could conceivably affect this dynamic between motivation and moral conscience in a variety of ways, depending heavily on context and the particular design features of the emotion in question.<sup>65</sup> For ex-

<sup>60</sup> See, e.g., Stanley Milgram, *Behavioral Study of Obedience*, 67 J. ABNORMAL & SOC. PSYCHOL. 371 (1963) (where participants administered electric shocks to a confederate at the direction of the scientist running the experiment); John M. Darley & Bibb Latané, *Bystander Intervention in Emergencies: Diffusion of Responsibility*, 8 J. PERSONALITY & SOC. PSYCHOL. 377, 379–81 (1968) (where participants did not aid a confederate faking a seizure when they believed other individuals were present).

<sup>61</sup> See, e.g., DALY & WILSON, *supra* note 59, at 128 (“Men are known by their fellows as ‘the sort who can be pushed around’ or ‘the sort who won’t take any shit,’ as people whose word means action and people who are full of hot air, as guys whose girlfriends you can chat up with impunity or guys you don’t want to mess with. In most social milieus, a man’s reputation depends in part upon the maintenance of a credible threat of violence.”). See generally RICHARD E. NISBETT & DOV COHEN, *CULTURE OF HONOR: THE PSYCHOLOGY OF VIOLENCE IN THE SOUTH* (1996).

<sup>62</sup> See Francesca Gino, *Understanding Ordinary Unethical Behavior: Why People Who Value Morality Act Immorally*, 3 CURRENT OPINION BEHAV. SCI. 107, 108–09 (2015) (covering studies in which individuals understand their behavior to be immoral and yet chose to engage when presented with the opportunity).

<sup>63</sup> See *id.*; see also DeScioli & Kurzban, *supra* note 50, at 286 (explaining that actor conscience competes with other motivations).

<sup>64</sup> See Patrick, *supra* note 56, at 1268–69 (discussing the phenomenon of two mental processes conflicting with each other and listing some common examples) (citing, among others, JONATHAN HAIDT, *THE HAPPINESS HYPOTHESIS: FINDING MODERN TRUTH IN ANCIENT WISDOM* 4–5 (2006) (“To understand most important ideas in psychology, you need to understand how the mind is divided into parts that sometimes conflict. We assume that there is one person in each body, but in some ways we are each more like a committee whose members have been thrown together to do a job, but who often find themselves working at cross purposes.”)). But see Albert Bandura et al., *Mechanisms of Moral Disengagement in the Exercise of Moral Agency*, 71 J. PERSONALITY & SOC. PSYCHOL. 364, 370–72 (1996) (discussing findings showing “moral disengagement”).

<sup>65</sup> See Nico H. Frijda, *The Laws of Emotion*, 43 AM. PSYCHOLOGIST 349, 349–50 (1988) (explaining emotions as responses to particular types of events); John Tooby & Leda Cosmides, *The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables*, in *HANDBOOK OF EMOTIONS* 114, 117–18 (Michel Lewis et al. eds., 3d

ample, some emotions are thought to be explicitly yoked to the outputs generated by the forecasting mechanisms of our moral conscience. These “moral emotions” such as shame, guilt, or embarrassment are thought to function specifically to move behavior *towards* our moral intuitions of right or wrong in cases where projected social devaluation would exceed other expected benefits.<sup>66</sup> One recent study, for instance, indicates that shame operates from an *ex ante* position to deter individuals from taking actions that would “cost” them more in terms of social devaluation than the other benefits the action would otherwise yield.<sup>67</sup> Another study found that the induction of guilt increased pro-social behavior *only when* it did not come at a personal cost to the participant.<sup>68</sup>

Other emotions, in other contexts, might function in the opposite manner, driving motivation *away from* judgments of moral acceptability when the overall expected benefits for doing so exceed expected social devaluation. For instance, emotions such as fear or disgust that function to help humans avoid potential threats might work to constrain motivation below the levels suggested by our moral conscience in instances where the risk of potential threats—physical or otherwise—reduce the expected payoffs of the action below the reputational value that would be gained by “doing the right thing.”<sup>69</sup> Consider, for example, the much-publicized murder of Kitty Genovese, a murder that was observed by numerous frightened witnesses who chose not to intervene or even to call the police.<sup>70</sup>

---

ed. 2008) (characterizing emotions are superordinate coordinating programs that evolved to respond to certain recurrent classes of events).

<sup>66</sup> See generally June Price Tangney et al., *Moral Emotions and Moral Behavior*, 58 ANN. REV. PSYCHOL. 345 (2007).

<sup>67</sup> Daniel Sznycer et al., *Shame Closely Tracks the Threat of Devaluation by Others, Even Across Cultures*, 113 PROC. NAT'L ACADEMY SCI. 2625, 2626 (2016) (“[S]hame is elicited by the prospect or actuality of negative information about the individual reaching others. Its neurocognitive architecture is designed to: (i) deter the individual from taking courses of action that would cost more in terms of social devaluation than the payoffs the action would otherwise yield . . .”).

<sup>68</sup> Ilona E. de Hooge et al., *What Is Moral About Guilt? Acting “Prosocially” at the Disadvantage of Others*, 100 J. PERSONALITY & SOC. PSYCHOL. 462, 467–71 (2011).

<sup>69</sup> See Tooby & Cosmides, *supra* note 65, at 117–18 (discussing the safety-prioritizing features of fear); see also Fessler et al., *supra* note 44, at 116 (finding that disgust decreases risk-taking in women); Lerner & Keltner, *supra* note 44, at 146 (explaining the finding that fearful people expressed pessimistic risk estimates and risk-averse choices).

<sup>70</sup> The precise details of the murder are disputed. As originally reported by the New York Times, “38 respectable, law-abiding citizens in Queens watched a killer stalk and stab a woman in three separate attacks in Kew Gardens . . . . Not one person telephoned – the police during the assault; one witness called after the woman was dead.” Martin Gansberg, *Who Saw Murder Didn't Call the Police*, N.Y. TIMES (Mar. 27, 1964), <http://www.nytimes.com/1964/03/27/37-who-saw-murder-didnt-call-the-police.html?mcubz=1> [https://perma.cc/W74G-2AZS]. Later accounts have insisted that this number was exaggerated, and that at least one witness shouted at the man to stop, and at least two called the police. However, at least one witness saw the attack, did nothing, and then took a nap rather than helping Genovese. Another opened a door, saw Genovese being stabbed, became scared, and closed the door. See generally, Nicholas Lemann, *A Call for Help: What the Kit-*

Heat of passion is concerned with a third possibility: that emotions could function to *increase* motivation beyond the level of acceptability generated by the moral conscience. Again, by hypothesis, many emotions could contribute to motivation exceeding moral conscience in contexts where the overall expected benefits from engaging in a particular act outweigh the expected social consequences. This is especially true of those emotions whose design features typically include increases in approach tendencies or behavioral motivation, like jealousy, love, sexual arousal, and, most saliently for heat of passion, anger.<sup>71</sup> While not directly supportive, the classic “heat of the moment” study by Ariely and Loewenstein is both illustrative and consistent with this theory.<sup>72</sup> In that study, male participants answered a series of questions about their willingness to engage in certain sexual behaviors, both in a sexually aroused and in an un-aroused state.<sup>73</sup> The study found that participants were significantly more likely to engage in many morally questionable behaviors (such as encouraging a date to drink or slipping her a drug to increase the chance that she would have sex) while in the aroused condition, suggesting that sexual arousal increased motivation beyond what the participants would do in a sexually un-aroused state.<sup>74</sup>

The specific assumption typically nested within the heat of passion framework is that anger pushes behavioral motivations beyond the level that the individual deems morally appropriate. Or, as it is commonly phrased, the individual proceeds not “from a bad or corrupt heart, but rather from the infirmity of passion to which even good men are subject.”<sup>75</sup> As previously outlined, while existing research has robustly demonstrated the motivational effects of anger, the relationship of this effect to judgments of moral acceptability—that is to say, the assumption at the heart of the doctrine—remains an open question.

Here, we tested this assumption in the context of the quintessential heat of passion case: a man witnessing or discovering the infidelity of his significant other.<sup>76</sup> Specifically, we tested whether, among males who had experienced the

---

ty *Genovese Story Really Means*, THE NEW YORKER (Mar. 10, 2014), <http://www.newyorker.com/magazine/2014/03/10/a-call-for-help> [https://perma.cc/N53V-ANQR] (outlining the disputed facts in the case and describing the facts as we now understand them).

<sup>71</sup> See Tooby & Cosmides, *supra* note 65, at 125 (discussing the different motivational properties of emotions and noting as an example that “a loss of face should increase the motivation to take advantage of opportunities for status advancement, and should decrease attention to attendant costs.”).

<sup>72</sup> Ariely & Loewenstein, *supra* note 51.

<sup>73</sup> *Id.* at 89–90.

<sup>74</sup> *Id.* at 91–95. But, again, because the study did not include measures of the participants’ moral judgments of their behavior, it is impossible to draw any concrete conclusions about this effect in relation to how acceptable participants believed those actions to be. *Cf.* Andrew Galperin et al., *Sexual Regret: Evidence for Evolved Sex Differences*, 42 ARCHIVES SEXUAL BEHAV. 1145 (2013) (finding that both men and women, but women moreso, tend to regret casual sexual encounters after the fact).

<sup>75</sup> Paz v. State, 777 So. 2d 983, 984 (Fla. Dist. Ct. App. 2000) (citation omitted).

<sup>76</sup> *Supra* notes 17–18 and accompanying text.

discovery of a partner's adultery, anger could increase motivations to retaliate beyond what the individual judged as morally acceptable.

### III. THE PRESENT STUDY

If the moral judgment/motivation dynamic at the core of the provocation doctrine is dependent on the various motivational, moral, and emotional mechanisms suggested by prior research, we should expect to observe the following psychological patterns: first, if moral conscience is a primary driver of behavioral motivations, then participants' judgments of moral acceptability should generally correlate with their behavioral motivations (*prediction #1*); second, to the extent that the correlation is not perfect, anger may have a greater effect on motivation than moral judgments (*prediction #2*), in which case anger should also account for variation in behavioral motivations beyond what can be accounted for by judgments of moral acceptability (*prediction #3*). However, this effect should not be uniform with respect to all behaviors: because anger functions to orchestrate aggression and punishment (or the threat of punishment) in order to increase the value placed on the angry individual's welfare, anger should motivate behavior beyond moral judgments of acceptability *only to the extent* that the behavior in question is salient to the function of anger (*prediction #4*).

#### A. Methods

##### 1. Participants

Our sample consisted of 1,966 adult males residing in the United States. Participants were recruited via Amazon's Mechanical Turk to complete an online survey administered through Qualtrics. Participants received monetary compensation for completing the survey, which was advertised as a "relationship survey" in which participants would be asked "a series of questions about a past relationship and about how you felt in that relationship." Of this sample, 1,236 (63 percent) met the participation criteria of having discovered the infidelity of significant other in a current or previous relationship. Twenty-seven participants were excluded after admitting to not meeting the participation criteria (after completing the survey and being assured of compensation regardless), yielding a total of 1,209 participants.

##### 2. Materials and Procedure

After passing the screening requirements, participants completed a series of items related to the circumstances surrounding their discovery of their partner's infidelity. Each participant answered two open-ended questions that required a minimum of 250-character responses. The first question asked participants to: *Please take 2–3 minutes and describe in a short paragraph the circumstances around the infidelity, what your partner did, and how you discovered it.* The



second question asked participants to: *Please take a moment to recollect how the harmful incident made you feel. Take 2–3 minutes and describe what you felt like in the moment you discovered your partner had been unfaithful.* Participants then reported how angry the discovery made the participants on a five point Likert-type<sup>77</sup> scale ranging from *not angry at all* to *extremely angry*, and completed two sections designed to measure both motivational and moral responses to the incident.

The motivation section presented participants with six retaliatory behaviors and asked them to recall and rate how much they wanted to engage in such behaviors on a five point Likert-type scale ranging from *not at all* to *extremely badly*. The behaviors consisted of three retaliatory behaviors salient to the function of anger in this context (i.e., to increase the value placed on the participant's welfare): (1) *yell at my partner*, (2) *push my partner*, (3) *hit or slap my partner*; and three behaviors not salient to the function of anger in this context (i.e. that might harm the partner but would not be expected to increase the value placed on the participant's welfare): (4) *steal money from my partner*, (5) *walk away from the relationship and never see her again*, and (6) *gossip about my partner to ruin her reputation*.<sup>78</sup> The moral section presented participants with the same behaviors and asked them to rate how justified they would have been to have engaged in such behaviors on a five point Likert-type scale ranging from *not justified at all* to *completely justified*.<sup>79</sup> All questions were randomized within sections. All analyses were completed in SPSS Version 24.

---

<sup>77</sup> A Likert-type scale is a common method in psychological research, usually used in surveys or questionnaires, where participants provide answers or ratings along a (typically) five or seven-point scale that varies by degree. For example, a participant might be asked how much they agree with a particular statement, and are given five options: (1) Strongly Disagree, (2) Disagree, (3) Neither Agree nor Disagree, (4) Agree, (5) Strongly Agree. *See generally* Gail M. Sullivan & Anthony R. Artino, Jr., *Editorial, Analyzing and Interpreting Data from Likert-Type Scales*, 5 J. GRADUATE MED. EDUC. 541, 541–42 (2013).

<sup>78</sup> Unlike yelling, pushing, and striking, the latter three behaviors (stealing money, gossiping, and walking away) do not involve the same direct aggression designed to alter the other individual's behavior. For example, stealing money and gossiping are passively aggressive, and involve taking action that is usually unbeknownst to the person towards whom they are directed. If the other individual does not know about the retaliation, their valuation on the actor's welfare cannot be recalibrated. In a similar vein, permanently walking away also forecloses the possibility of recalibration. Compounding this effect is that, given the dynamic of the participants (i.e. male towards female), physical violence and threats of physical violence carry a greater chance of effectiveness and a reduced chance of retaliation, due to the average difference in strength and physical stature. *See* Sell et al., *supra* note 39, at 15073–78 (explaining the logic of these differences and showing this dichotomy in a study comparing the relationship among strength, anger, and aggression between men and women). For an interesting discussion of how wife-killing can be driven by a process designed to recalibrate welfare valuations, *see* Joshua D. Duntley & David M. Buss, *The Plausibility of Adaptations for Homicide*, in *THE INNATE MIND: STRUCTURE AND CONTENTS* (Peter Carruthers et al. eds., 2005).

<sup>79</sup> Moral acceptability is operationalized here as participant evaluation of how justified their action would have been. *Compare* Justify, “to show to be just, right, or in accord with reason.” WEBSTER NEW COLLEGE DICTIONARY, THIRD EDITION 734 (2005), with Moral, “good

### B. Analysis and Results

To assess the various associations among the variables, we conducted a series of correlational analyses between (1) motivation and moral judgments, (2) anger and motivation, and (3) anger and moral judgments. We also conducted partial correlation analyses between (4) anger and motivation while controlling for moral judgments, and (5) anger and moral judgments while controlling for motivation. Figure 1 presents these five correlational relationships for each of the six behaviors.

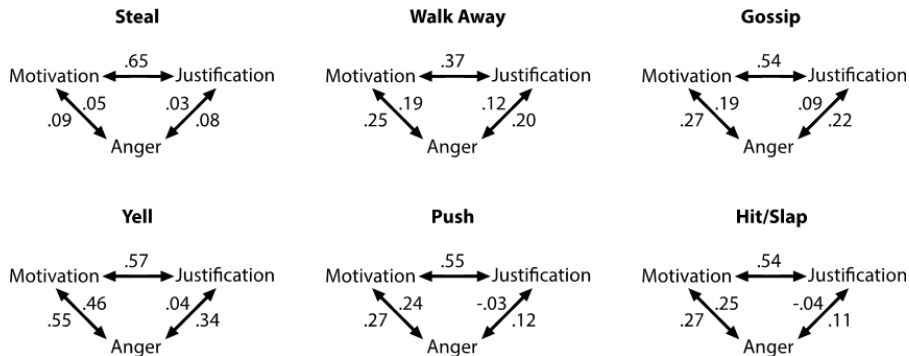


FIGURE 1. CORRELATIONS AMONG ANGER, MOTIVATION, AND MORAL JUDGMENTS FOR EACH BEHAVIOR. ZERO-ORDER CORRELATIONS ARE REPORTED OUTSIDE OF THE LINES, PARTIAL CORRELATIONS ARE REPORTED INSIDE OF LINES. ALL CORRELATIONS ABOVE .05 WERE SIGNIFICANT AT A MINIMUM LEVEL OF  $p < .01$ . ALL CORRELATIONS .05 OR LOWER WERE NOT SIGNIFICANT.

As predicted, there were significant correlations between motivation and moral judgment for all six behaviors. Moreover, anger ( $M = 4.11$ ,  $SD = 1.02$ ) was significantly positively correlated (zero-order) with all six motivational items, as well as all six moral acceptability items. However,  $z$ -tests designed to measure the relative correlations for each behavior revealed that the correlations between anger and motivation were significantly stronger than the correlations between anger and moral acceptability for: yell  $z = 9.17$ ,  $p < .001$ , push  $z = 5.31$ ,  $p < .001$ , hit or slap  $z = 5.96$ ,  $p < .001$ , but not steal money  $z = .42$ ,  $p = .677$ , walk away  $z = 1.60$ ,  $p = .109$ , gossip  $z = 1.88$ ,  $p = .059$  (all two-tailed).<sup>80</sup> In other words, while higher levels of anger were associated with both

or right in conduct or character;" WEBSTER NEW COLLEGE DICTIONARY, THIRD EDITION 882 (2005).

<sup>80</sup> See Ihno A. Lee & Kristopher J. Preacher, *Calculation for the Test of the Difference Between Two Dependent Correlations with One Variable in Common* (Sept. 2013), <http://quantpsy.org/corrtest/corrtest2.htm> [<https://perma.cc/3XYA-HUXX>] (for computer software that executes the test for comparing correlations sharing a variable); see also James H. Steiger, *Tests for Comparing Elements of a Correlation Matrix*, 87 PSYCHOL. BULL. 245, 245-47 (1980) (explaining the theoretical underpinnings of the test).

increased motivation *and* increased judgments of moral acceptability, the association with anger was significantly *stronger* with motivation only for those behaviors salient to anger's function (see Figure 2).

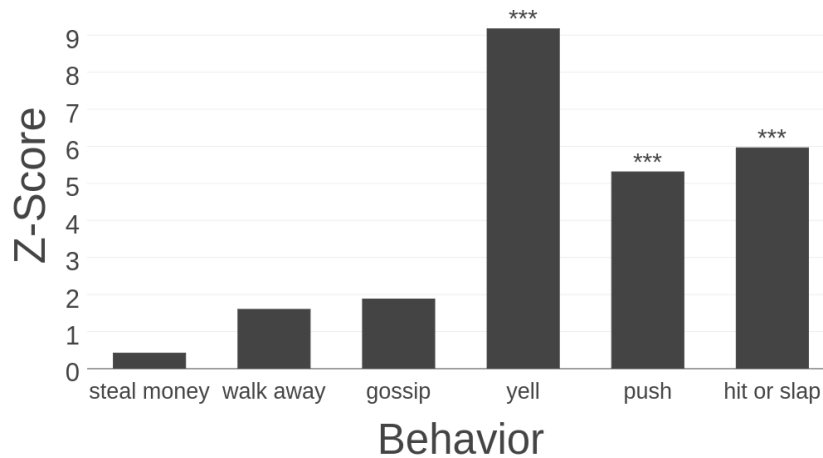


FIGURE 2. Z-SCORES (TWO-TAILED) OF THE DIFFERENCES BETWEEN THE CORRELATION OF ANGER WITH MOTIVATION AND THE CORRELATION OF ANGER WITH JUDGMENTS OF MORAL ACCEPTABILITY. \*\*\*  $P < .001$ .

Next, when controlling for moral acceptability, anger remained significantly positively correlated with motivation for all behaviors except stealing money. Conversely, while controlling for motivation, anger remained significantly positively correlated with judgments of moral acceptability for walk away and gossip, but not for steal money, yell, push, or hit or slap. Thus, most saliently for the heat of passion doctrine, and as shown both in Figure 1 and in Figure 3, the relationship between anger and participants' judgments of moral acceptability for yelling, pushing, and hitting or slapping their partner could be entirely accounted for by anger's association with motivation, but the reverse was not true: anger predicted motivation to yell and engage in physical violence beyond what could be predicted by judgments of moral acceptability.

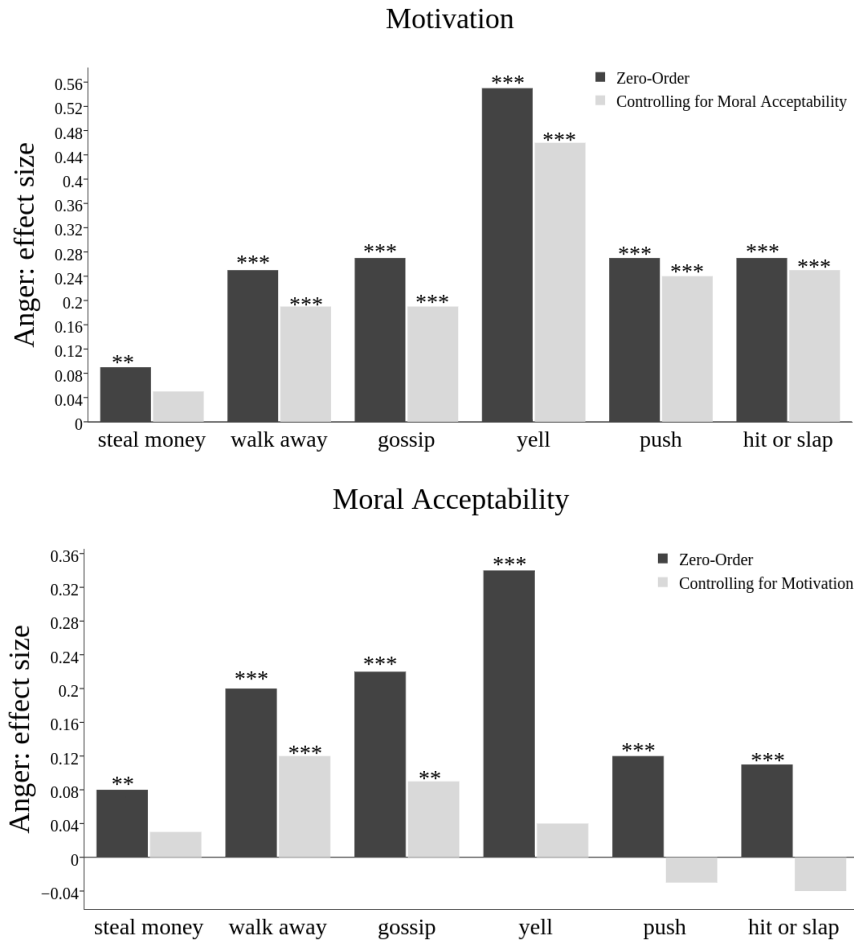


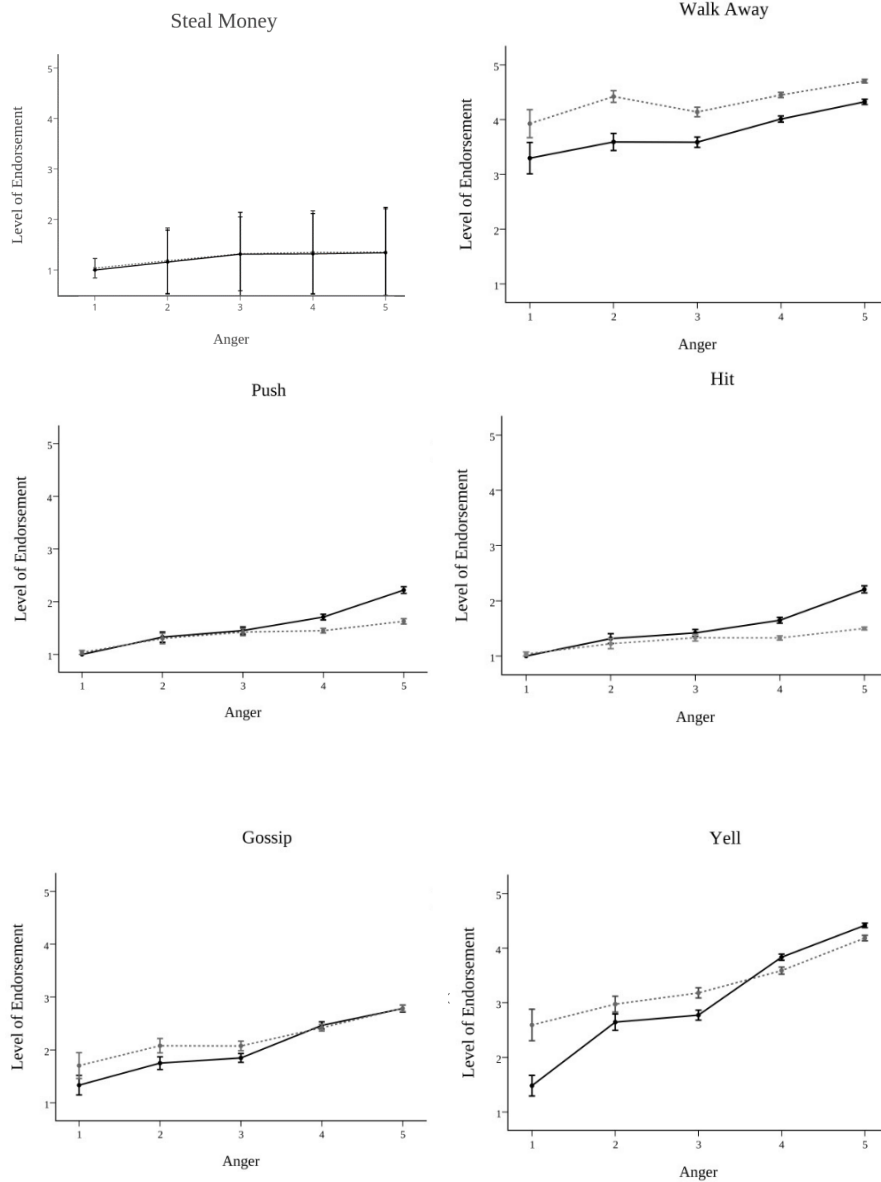
FIGURE 3. CORRELATION EFFECT SIZES BETWEEN ANGER AND MOTIVATION, AND ANGER AND MORAL ACCEPTABILITY, EACH AT A ZERO-ORDER LEVEL AND WHILE CONTROLLING FOR THE OTHER VARIABLE. \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Because motivation and moral acceptability were rated on different Likert-type scales, we cannot draw any concrete conclusions by comparing their respective means. However, the various effects demonstrated via our correlational analyses at least *appear* to be corroborated by the relative change of means for motivation and moral judgments at the different levels of anger. As Figure 4 illustrates, the relationship among the three measures varies depending on the behavior in question.

FIGURE 4. MEAN LEVELS OF REPORTED MOTIVATION AND JUDGMENTS OF MORAL ACCEPTABILITY.

ERROR BARS REPRESENT ONE STANDARD DEVIATION.

—●— Motivation  
 - - ● - - Moral Acceptability



For stealing money, the ratings of motivation and moral acceptability are nearly identical on their respective scales at every level, despite the very large standard deviations in answers. For walking away, judgments of moral acceptability are higher than motivation on their respective scales at every stage of anger. By comparison, with respect to yelling or gossiping, at low levels of anger, participants had higher ratings for moral acceptability than motivation, but at high levels of anger, these two judgments collapse or, in the case of yelling, reverse—so that motivation ratings now exceed moral judgments on their respective scales. The most salient behaviors for the heat of passion doctrine—the two physical violence measures—have a similar pattern. At low levels of anger, the ratings are nearly identical (both extremely low) on their respective scales. But at higher levels of anger, the two begin to separate, with motivation ratings exceeding moral acceptability ratings on their respective scales.

#### IV. DISCUSSION

Four predictions regarding the psychology of provocation derived from the existing theoretical literature were either fully or partially supported. These findings provide evidence that (a) moral judgments and motivation are generally correlated, (b) anger has a stronger association with motivation to engage in certain retaliatory behaviors than with moral judgments as to how justified such behaviors would be, (c) anger can account for variation in motivation beyond what can be accounted for by judgments of moral acceptability, and (d) that these effects are specific to those behaviors that are salient to the function of anger. Taken together, these findings lend support to a longstanding psychological assumption of human behavior in the law—that individuals in a heightened state of anger may behave in ways that they do not otherwise feel are morally acceptable.

That support, however, is accompanied by an extensive set of caveats. While our results support the general logic of the doctrine, they also call other more specific aspects, and particular variations, of the rule into question. Moreover, there are methodological limitations inherent in our approach, and a risk of extracting erroneous normative conclusions from purely descriptive tests. We discuss these limitations, as well as other normative implications, below.

##### A. *Study Limitations*

Like all empirical studies, ours has its own set of methodological limitations. Among them are the following four.

First, the retrospective nature of the study means that participants are being asked to recall how they felt during an event that may have occurred years or even decades ago. The validity of the participants' responses depends on their

ability to accurately recall these events and feelings—an ability that can be attenuated by a number of things, including both time and intervening events.<sup>81</sup>

Second, the study was designed to test the psychological mechanisms that underlie behavioral decisions made ‘in the heat passion,’ but, out of fear that participants would not respond honestly, we did not obtain measures of (a) what participants actually *did*, as opposed to what they were motivated to do, or (b) whether participants were motivated to actually kill their unfaithful partner. Thus, conclusions as to whether these results extend to either of those conditions require an additional level of inference.<sup>82</sup>

Third, there are well-documented problems with relying on Amazon Mechanical Turk (“MTurk”) samples in psychological research. MTurk “workers” who complete surveys for money tend to be, compared to the general population: younger, overeducated, underemployed, less religious, and more liberal, with Asians being overrepresented and Blacks and Hispanics underrepresented.<sup>83</sup> Moreover, many MTurk workers tend to be “expert” participants, who become familiarized with many of the common methodologies used for psychological research, suggesting that their responses might be different from the traditionally “naïve” subjects.<sup>84</sup> These factors cast doubt on what researchers call the “external validity”—or the ability to generalize to the population at large—of inferences made using MTurk samples.

Finally, the study is not experimental and so we are limited in our ability to draw conclusions about causality. While we are able to conclude that the association of anger is stronger with motivation than with moral judgments, and that anger can explain variation in motivation that cannot be accounted for by how morally acceptable participants believed those actions to be, we are unable to definitely conclude that anger *causes* this discrepancy. That being said, it would be extremely difficult from a practical perspective—and precarious from an ethical perspective—to recreate the conditions of a typical provocation sce-

---

<sup>81</sup> See generally Elizabeth Loftus, *Our Changeable Memories: Legal and Practical Implications*, 4 NATURE REVIEWS: NEUROSCIENCE 231 (2003). But see Dorte Berntsen & David C. Rubin, *Emotionally Charged Autobiographical Memories Across the Life Span: The Recall of Happy, Sad, Traumatic, and Involuntary Memories*, 17 PSYCHOL. & AGING 636, 642–46 (2002) (showing that emotional experiences are remembered more vividly and forgotten more slowly than neutral events).

<sup>82</sup> But see *infra* Section IV(B).

<sup>83</sup> Gabriele Paolacci & Jesse Chandler, *Inside the Turk: Understanding Mechanical Turk as a Participant Pool*, 23 CURRENT DIRECTIONS PSYCHOL. SCI. 184, 185 (2013) (“In general, workers are diverse but not representative of the populations they are drawn from, reflecting that the Internet users differ systematically from non-Internet users. Workers tend to be younger (about 30 years old), overeducated, underemployed, less religious, and more liberal than the general population. Within the United States, Asians are overrepresented and Blacks and Hispanics are underrepresented relative to the population as a whole.”) (citations omitted).

<sup>84</sup> See Jesse Chandler et al., *Using Nonnaive Participants Can Reduce Effect Sizes*, 26 PSYCHOL. SCI. 1131, 1137–38 (2015) (summarizing the effects of exposure to research materials on effect sizes).

nario and induce the levels of anger necessary for study in a lab. Instead, we accepted the limitations inherent in using recall to study a sample of individuals that hew closely to the archetypal real-world case.

### B. Normative Implications

The enterprise of the behavioral sciences is ultimately a descriptive one—to explain why and how humans behave in the ways that they do. The enterprise of the law, on the other hand, is driven by normative objectives: to choose a socially desirable state of affairs and push human behavior towards that state of affairs. In the end, any behavioral insights, including our results here, can only inform such normative questions, not decide them. Thus, while there are normative lessons to be gleaned from these results, we suggest caution with deriving any immediate prescriptions.

Our study supports the idea that angry individuals may behave in certain ways that they feel are morally unacceptable, and ways that individuals in calm and collected states would not. Another way to state or interpret these results is that, at low levels of anger, judgments as to how justified or morally acceptable a behavior is carry more weight in determining motivation to act than they do at high levels of anger. That being said, our results here do nothing to settle why—or whether—acting in the heat of passion *should* mitigate punishment.<sup>85</sup> Our results do not answer, for example, the complicated questions that revolve around whether the capacity to act in the heat of passion is a less “blameworthy” character flaw than calculated bad intent, as some courts and scholars posit,<sup>86</sup> or whether such individuals are less deserving of punishment. These policy questions are part of a separate enterprise.<sup>87</sup>

While such questions remain open, our findings here do add some clarity to various components of the calculation. Take, for example, the question of whether the individual is acting “without reason.” Our results make clear that there is a definite logic to the provoked actor’s behavior, even where that logic is not consciously accessible. Anger, at high levels, appears to override judgments of moral acceptability, and does so in consistent, predictable ways. Anger, for example, does not influence one’s motivation to walk away from a relationship or to spread nasty gossip about one’s partner over and above its influence on judgments of how acceptable it would be to do so. It does, however, coincide with a motivation to yell, push, or to hit one’s partner beyond what judgments of moral acceptability would dictate. In other words, anger does not

<sup>85</sup> See *supra* notes 31–36 and accompanying text.

<sup>86</sup> See *e.g.*, notes 33 and 35.

<sup>87</sup> See Patrick, *supra* note 56, at 1286 (“By deconstructing the functions and operations of moral intuitions . . . we avoid being wooed by their subconscious origins and falling prey to the naturalistic fallacy. That is, we avoid making the assumption that what is natural is per se good, and avoid committing the logical fallacy of assuming *ought* from *is*.”) (citing Owen D. Jones & Timothy H. Goldsmith, *Law and Behavioral Biology*, 105 COLUM. L. REV. 405, 485 n.234 (2005)).



flip a switch that results in random, unpredictable behavior; anger reliably entrains one's motivational systems in a way that is consistent with its ultimate function—to increase the value others place on one's welfare. This is stated perhaps most succinctly by Kyron Huigens, who notes that

The insane actor acts *against* reason, but the provoked actor does not. Very much to the contrary: the provoked actor acts *for* reasons. The paradigmatic cuckolded husband acts as he does for the reason that he has been cuckolded. He might not act with cool deliberation, but he does act for a reason.<sup>88</sup>

In a similar vein, our results also question the conclusion that acts committed in the heat of passion are somehow less deterrable than acts that are made with calculated deliberation. Instead, our results are consistent with the theory that the various moral, emotional, and motivational systems at play in a typical heat of passion scenario are actually engaged in a sophisticated cost-benefit analysis that weighs the potential benefits of certain behaviors against the potential negative consequences of engaging in them.<sup>89</sup> And while further research is needed to hone in on the exact parameters of these systems and the degree to which the threat of punishment can influence our moral judgments and behavioral motivations, it is reasonable to conclude that the threats of punishment and social condemnation (often through our moral conscience) have an effect on our motivations to act.<sup>90</sup> After all, for every one individual who kills in the heat passion, there are scores of others who might be similarly enraged, but refrain from acting. This sentiment is vividly demonstrated by participants' answers to our prompt to *describe what you felt like in the moment you discovered your partner's infidelity*, which included, for example:

I felt like killing someone. I swear I could've lifted a car;  
I was ready to kill both of them;  
I still do not know how I managed to not kill her;  
It was probably the one time in my life that I could have committed double homicide and not had any remorse;  
I wanted her and him dead. I just wanted both of them removed from life;  
I was so angry that had they both or either one of them been around, I may well have ended up in prison for murder;  
I wanted her dead. Extreme anger and rage. I wanted to strangle the life out of her. She was a heartless bitch; and  
I felt like beating her to death. I wanted her to die, and I wanted the last thought to go through her head to be that she brought this on herself. If I could have done it without getting caught, I believe I would have killed her and not felt any remorse.

---

<sup>88</sup> Kyron Huigens, *A Critical Introduction to the Symposium*, 43 U. MICH. J.L. REFORM 1, 7 (2009) (paraphrasing John Gardner's argument and citing generally JOHN GARDNER, OFFENCES AND DEFENCES: SELECTED ESSAYS IN PHILOSOPHY OF CRIMINAL LAW (2007)).

<sup>89</sup> See *supra* notes 52–55 and accompanying text.

<sup>90</sup> See *supra* notes 52–55 and accompanying text; see generally Max M. Krasnow et al., *What Are Punishment and Reputation for?*, 7 PLOS ONE 1 (2012).

These statements raise another important normative limitation of our findings: they do nothing to allay the concerns that the heat of passion defense is often used to mitigate the punishment for male violence against women. A prominent criticism of the doctrine is that it is an “abuse excuse” and that it has been disproportionately beneficial to cuckolded males while leaving other categories of defendants—such as females who kill their husbands after years of physical abuse, but do so not while “in the heat of passion”—without a similar safety net.<sup>91</sup> These criticisms are unaffected by the results reported here. Again, what our findings support is the idea that the general psychological premises on which these normative conclusions are based are, to a degree, accurate. They do not support (nor do they contradict) the underlying normative claim that any particular class of defendants is any less deserving of punishment than others.

Finally, our findings should caution against some of the more liberal standards being deployed in the wake of the MPC’s expanded definition.<sup>92</sup> Emotions are not a unitary phenomenon. Just as anger predictably influenced motivation in ways that were consistent with its psychological function, we should expect other emotions to have their own set of effects unique to their individual functions. Anger has a much different function than fear, which has a much different function than disgust, and so on. There is no *a priori* reason, based on our findings here, to assume that because anger coincides with motivation beyond what the actor feels is justified, that in turn any “extreme mental or emotional disturbance” could produce this same effect. In fact, existing theory and research would suggest an alternative hypothesis—that only those emotions whose design features typically include “approach” motivations might reproduce the effect seen with anger.<sup>93</sup> More research would be necessary to substantiate this claim.

#### CONCLUSION

These findings are the first in any discipline to empirically demonstrate what, for centuries, the criminal law has simply assumed. Do these results con-

---

<sup>91</sup> See Carolyn B. Ramsey, *Provoking Change: Comparative Insights on Feminist Homicide Law Reform*, 100 J. CRIM. L. & CRIMINOLOGY 33, 33 (2010) (discussing the concept of provocation as an “‘abuse excuse’”); The critique is so prominent that it appears alongside the doctrine itself in leading criminal law casebooks. Gruber, *supra* note 19, at 273 (noting the prevalence of the critique). For in-depth examples of this critique, see generally Clavel, *supra* note 27; Coker, *supra* note 15; Miller, *supra* note 18; and Rozelle, *supra* note 18.

<sup>92</sup> See, e.g., OR. REV. STAT. § 163.118 (2017) (“Criminal homicide constitutes manslaughter in the first degree when: . . . (b) It is committed intentionally by a defendant under the influence of extreme emotional disturbance as provided in ORS 163.135, which constitutes a mitigating circumstance reducing the homicide that would otherwise be murder to manslaughter in the first degree and need not be proved in any prosecution.”); UTAH CODE ANN. § 76-5-205.5 (1)(b) (2009) (“Special mitigation exists when the actor causes the death of another or attempts to cause the death of another: . . . (b) under the influence of extreme emotional distress for which there is a reasonable explanation or excuse.”).

<sup>93</sup> See *supra* notes 71–74 and accompanying text.

Fall 2017]

*NOT FROM A WICKED HEART*

59

firm that the heat of passion doctrine has had it right all along? No, and in fact they suggest that, to some degree, there are flaws in the normative reasoning that has traditionally been applied. But they do allow courts and legal scholars to engage in future normative analysis with a clearer and more confident understanding of the psychological processes on which those normative analyses depend.

These results should not be seen as the definitive empirical treatment of either the provocation doctrine or the psychological phenomena on which the doctrine relies. Rather, this study should serve as a first step—as proof of concept for a general framework that triangulates the relationships among emotions, moral judgments, and motivation.

[PAGE INTENTIONALLY LEFT BLANK]