PROVOCATION, PERCEPTION, PASSION

Stephen P. Garvey

Abstract

A theory of the provocation doctrine—labeled the "partial forfeiture" theory—is developed based on a particular understanding of the psychological process by which a defendant "loses self-control." The interpretation this psychology would assign to each of the provocation doctrine's elements is explained. According to this theory, when a defendant could not, due to passion, have done otherwise than form an intent to kill, he's nonetheless guilty of murder if the state regards him as someone who does not habitually *see or perceive* the moral world in the way the state obligates it to be seen or perceived. The theory thus portrays the state, when state actors apply the provocation doctrine to the facts of a particular case, as making a judgment, not about the way in which the defendant judges the "moral world," but instead about the way in which he's disposed to see or perceive it.

About the Author

A. Robert Noll Professor of Law, Cornell Law School. I thank Joshua Dressler for comments on an earlier draft.

Table of Contents

Introduction

When the elements constituting the doctrine of provocation, or heat of passion, have been established, a defendant who would otherwise have been liable for "murder" is instead liable only for the lesser offense of "manslaughter." Academics who write about the provocation doctrine generally fall into one of two groups. Writers in the first group believe some particular formulation of the doctrine, found in some jurisdiction or another, should be changed in some way: they hope to persuade someone, presumably some state authority, to alter or re-interpret some particular formulation of the doctrine, to achieve some end or another.

Writers in the second group, in contrast, do not imagine themselves to be advocating for any change in the doctrine. Instead, they hope to persuade someone, presumably anyone inclined to listen, to believe as they do: that the doctrine, stated to some extent in abstraction from any particular formulation, mitigates murder to manslaughter because the doctrine is one kind of thing (e.g., a "partial excuse") rather than another (e.g., a "partial justification"). These academics aim to give a "theory" or "rationalization" of the defense, not as it ought to be, but as it is currently instantiated in positive law. My aim here is to contribute to the work of this second group: to offer another "theory" or "rationalization" of the provocation doctrine as it exists in positive law, not to offer reasons to change or reform the doctrine.

Writers in the first group have made a range of reform proposals, with which other writers almost invariably disagree; for example:

- Some writers condemn the Model Penal Code's (MPC) formulation of the doctrine: it leads, they say, to a "murder law that is both illiberal and often perverse."[1] Others see things differently. Coming to the Code's defense, they say the MPC formulation "deserves considerable attention," even if it should be modified here or there.[2]

1. *See, e.g.*, Victoria Nourse, *Passion's Progress: Modern Law Reform and the Provocation Defense*, 106 Yale L.J. 1331, 1332 (1997).
2. *See, e.g.*, Joshua Dressler, *Why Keep the Provocation Defense?: Some Reflections*

- Some urge that a necessary, but apparently not sufficient, condition for finding an alleged provocation to be "adequate" should be that the decedent's action "reflect[] a wrong that the law would independently punish."[3] Others prefer a broader definition of "adequacy," one designed to leave the question more or less in the hands of a jury, asking its members to decide if, for example, the alleged provocation "might cause a reasonable or ordinary person of average disposition, to lose self-control and act rashly and without due deliberation."[4]
- Some say the doctrine's elements should be pruned; for example, the requirement of adequate provocation, they say, has nothing to do with a defendant's culpability, and should thus be excised from the doctrine altogether.[5] Others demur: removing the need for a defendant to show that the provocation to which he responded was adequate would have "significant disadvantages."[6]
- Some say that particular kinds of alleged provocations, like actions described as "nonviolent homosexual advance," should be defined by statute to be inadequate as a matter of law.[7] Permitting jurors even to consider such an alleged provocation amounts, they say, to an "institutionalization of homophobia."[8]

---

on a Difficult Subject, 86 Minn. L. Rev. 959, 984 (2002); *Id.* at 989 (concluding that the "EMED [extreme mental or emotional disturbance] formula crafted by the American Law Institute is flawed, but includes many positive features").

3. *See* Nourse, *supra* note 1, at 1396. Notwithstanding this simple test for "adequacy," Nourse goes on in a footnote to say that deciding if an alleged provocation constitutes "adequate" provocation will, at least sometimes, require a factfinder to adjudicate between, or somehow balance, competing and conflicting "norms." *Id.* at 1396 n.386. In light of these caveats, one commentator suggests that "Nourse's seemingly bright-line warranted excuse/legality rule looks more akin to the MPC's standard of 'reasonable explanation or excuse,'" which is the standard Nourse explicitly rejects. *E.g.*, Aya Gruber, *A Provocative Defense*, 103 Calif. L. Rev. 273, 290 (2015).
4. *See* Dressler, *supra* note 2, at 998–99.
5. *See, e.g.*, Richard Singer, *The Resurgence of Mens Rea: I – Provocation, Emotional Disturbance, and the Model Penal Code*, 27 B.C. L. Rev. 243, 315 (1986) ("In assessing . . . liability, the jury should consider only the defendant and his emotional characteristics; no suggestions of "reasonable" people or "adequate" provocation should be allowed.").
6. *See, e.g.*, A.J. Ashworth, *The Doctrine of Provocation*, 35 Cambridge L.J. 292, 319 (1976) ("[T]he abolition of the objective condition [i.e., "adequate" provocation] can be seen to have significant disadvantages in many respects.").
7. *See, e.g.*, Robert B. Mison, Comment, *Homophobia in Manslaughter: The Homosexual Advance as Insufficient Provocation*, 80 Calif. L. Rev. 133, 176–77 (1992) ("[J]udges should . . . find as a matter of law that a homosexual advance is insufficient provocation."). Some states, following the advice of the American Bar Association, have in fact made "non-violent homosexual advance," defined in one way or another, "inadequate" as a matter of law. Another proposal, along similar lines, is to remove some provocations, such as the sight of adultery, from among those provocations defined to be "adequate" as a matter of law.
8. *Id.* at 136.

Others disagree: "[A] special rule precluding the use of the provocation defense in homosexual advance (or, more generally, sexual advance) cases is too tenuous to withstand scrutiny."[9]

- At the extreme, some say the doctrine itself is a relic, a "product of a bygone view regarding male honor,"[10] and should be abolished altogether. Its continued existence simply "reinforces in the law that which public institutions ought in fact to be seeking to eradicate."[11] Others believe the case for abolition has not yet been made. If the provocation doctrine identifies, from among the class of all defendants who intend to kill, a subclass who is less culpable than the rest, then "abolitionist arguments are inadequate to the task."[12]

Writers in the second group, unlike those above, do not try to give reasons to change or reform existing law. Instead, they give reasons for classifying or understanding the defense in one way rather than another.

9. *See, e.g.*, 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, 729 (1995).

10. Gruber, *supra* note 3, at 286.

11. *See, e.g.*, Jeremy Horder, Provocation and Responsibility 194 (1992) ("From a feminist perspective the existence of [the provocation doctrine] simply reinforces in the law that which public institutions ought in fact to be seeking to eradicate"). For Horder, "[i]t is one thing to feel great anger at great provocation; but quite another (ethical) thing to experience and express that anger in [physically] retaliatory form," *id.* at 197, which leads him to conclude that no provocation can qualify as "adequate," and thus that the doctrine should for that reason be abolished.

12. *See, e.g.*, Dressler, *supra* note 2, at 962–63 ("[A]bolitionist claims are inadequate to the task [of abolishing the defense] as long as the defense is understood in just-deserts-loss-of-self-control terms"). The "abolitionist claims" Dressler believes are "inadequate" are claims said to follow from consequentialist arguments. A consequentialist argument in favor of abolition concedes that a defendant who forms an intent to kill, but who also establishes the provocation doctrine's elements, is somehow "less culpable" for having formed that intent compared to a defendant who forms an intent to kill but who fails to establish the doctrine's elements. Nonetheless, so the abolitionist argument goes, the former defendant, like the latter, should be convicted of murder, despite being less culpable. That's because the proponent believes doing so will have consequences the proponent believes are good consequences; moreover, these good consequences are offered as reasons to "justify" convicting the less culpable defendant of the same crime (murder) as the more culpable defendant. In short, the proponent believes some anticipated good justifies an admittedly disproportionate punishment.

Of course, other writers who evaluate the doctrine based on the consequences predicted to result from its continued existence identify and balance differently the consequences thought to be relevant. *See, e.g.*, Gruber, *supra* note 3, at 273 (The "feminist critique of provocation . . . may unintentionally instantiate and entrench punitive impulses that create and sustain mass incarceration," i.e., "entrench[ing] punitive impulses" and "sustain[ing] mass incarceration" are bad consequence predicted to result from abolishing the doctrine, which should be balanced against whatever good consequences are predicted to result from its abolition.). Consequentialist reasoning predictably produces disagreements of this kind.

In so doing, they aim to explain why the doctrine has the mitigating effect it has and thus what the state, acting through a jury, is doing when a jury finds that a defendant has established the doctrine's elements and thus convicts the defendant of manslaughter rather than murder. In short, they aim to give a "theory" of the provocation doctrine. Helpfully or unhelpfully, the language of "excuse" and "justification" is the language most writers use to characterize why it is they believe the doctrine has the mitigating effect it has, and thus what type of theory they are offering.[13] For example:

- Some writers say the provocation doctrine is a "partial excuse," such that any defendant who establishes its elements is partially excused for having formed an intent to kill and is thus liable for manslaughter rather than murder.[14]
- Others say that the provocation doctrine is a "partial justification,"[15] such that any defendant who establishes its elements is partially justified in having formed an intent to kill, in the sense that establishing the doctrine's elements mean the defendant is for some reason guilty of a lesser wrong compared to a defendant who fails to establish the doctrine's elements.
- Still, others say that the provocation defense somehow combines both excuse and justification,[16] such that any defendant who

13. The concepts of "excuse" and "justification" can be rendered or understood from the "point of view" of something commonly called "morality," or from the "point of view" of the "law," limited in the present context to the "criminal law." This, I take it, is the basic thesis argued for in Mitchell N. Berman, *Justification and Excuse, Law and Morality*, 53 Duke L.J. 1 (2003). I am partial to Berman's thesis, which, if true, means that when the state describes a criminal defendant's action as "justified," all it's "really" saying is that the defendant's action is not "criminal," and when the state describes a criminal defendant's action as "excused," all it's "really" saying is that the defendant's action is not "punishable." Having said that, the text will nonetheless speak, consistent with prevailing convention, in terms of excuse and justification (where the latter is understood as equivalent to permission).

14. Joshua Dressler's work, beginning with Joshua Dressler, *Rethinking Heat of Passion: A Defense in Search of a Rationale*, 73 J. Crim. L. & Criminology 421 (1982), and spanning decades thereafter, is most prominently associated with this theory.

15. The usual citation is to Finbarr McAuley, *Anticipating the Past: The Defence of Provocation in Irish Law*, 50 Mod. L. Rev. 133 (1987). Indeed, McAuley states point blank: "[T]he defence of provocation functions as a partial justification rather than a partial excuse." *Id.* at 139. On my reading, McAuley's understanding of the provocation doctrine is much the same as that of Kahan, Nussbaum, and Nourse, all of which seems to me to assume a psychology in which the defendant's reason was "disturbed" by "hot blood," but not one in which passion "incapacitated" the defendant's reason to any degree, and thus not one in which the defendant "lost self-control." *See infra* note 34.

16. *See, e.g.*, Mitchell N. Berman & Ian P. Farrell, *Provocation Manslaughter as Partial Justification and Partial Excuse*, 52 Wm. & Mary L. Rev. 1027, 1101 (2011) ("Our account [of the provocation defense] . . . would require that the defendant be in the heat of passion *and* that the provocation be of the sort that gives the

establishes its elements is, in some sense and at the same time, both justified and excused for having formed an intent to kill.

What's thought to be at stake in this theoretical debate is what one understands the state to be doing, if and when it makes the provocation doctrine a part of its criminal law. Is it trying to identify conditions, reflected in the doctrine, under which a defendant who forms an intent to kill (and who typically does kill) has done something less wrongful compared to an otherwise similarly situated defendant? That would be to portray the doctrine as a partial justification. Or, is it trying to identify conditions under which a defendant who forms an intent to kill is less culpable for having formed that intent compared to an otherwise similarly situated defendant? That would be to portray the doctrine as a partial excuse.

What follows offers yet another such "theory" of provocation. Accordingly, its aim is not to praise or condemn the provocation doctrine, but rather to better understand why it reduces murder to manslaughter when its elements have been established. I would not try to pigeonhole this theory, dubbed here the partial forfeiture theory, into the conventional partial justification-partial excuse dichotomy. The theory is what it is, however it might be characterized.

The partial forfeiture theory resembles, in its general structure, a better-known theory, commonly called in the literature the partial excuse theory. The partial excuse theory is also set forth below, mainly for purposes of comparing and contrasting it to the partial forfeiture theory. Both theories are structurally similar, but the partial forfeiture theory

---

defendant reasons *to kill*, subject to a substantial qualification") (emphasis added). Berman and Farrell believe that Andrew Ashworth has likewise portrayed the provocation doctrine as combining a "justification" element and an "excuse" element in the way they propose. *See* Berman & Farrell, *id.* at 1058 (citing A.J. Ashworth, *The Doctrine of Provocation*, 35 Cambridge L.J. 292 (1976)). *But see* Berman & Farrell, *id.* at 1058 n.119 (noting that Ashworth's account of the provocation doctrine can also fairly be read to combine an excusing condition with a forfeiture condition).

In a previous effort, I tried to combine provocation's elements in a different way, according to which the doctrine required proof of "adequate provocation," not because that element served any purpose independent of the "loss of self-control" element, but because proof of "adequate provocation" provided evidence necessary to establish that the defendant in fact "lost self-control." *See* Stephen P. Garvey, *Passion's Puzzle*, 90 Iowa L. Rev. 1677, 1733 (2005). I no longer believe that effort was successful, at least not as a "theory" of the provocation doctrine as it exists in positive law, as opposed to a proposal for another partial defense of some sort. *But see infra* note 41. Among other things, it mistakenly construed the "loss of self-control" element as equivalent to what's described *infra* as "abandoning self-control." *See infra* pt. II.B. However, the provocation doctrine, as I now understand it, or at least as I now present it, requires a defendant to have "lost self-control," where "losing self-control" is different from "abandoning" it. In short, I misunderstood what it means to say of a defendant, for purposes of the provocation doctrine, that he "lost self-control." Berman and Farrell fairly criticize this prior effort at Berman & Farrell, *supra* at 1061–62.

differs, one hopes, in interesting ways from the partial excuse theory. Both theories are plausible theories of the provocation doctrine in the sense that each offers a coherent account of the doctrine's elements. What makes the two theories similar, and what makes them different, will take time to develop, but here's a preview.

The provocation doctrine, as I see it, has two broad parts or features or conditions: an *excusing condition* and a *forfeiture condition*.[17] Combined in some way, they produce a "partial defense."[18] The partial excuse

17. "Forfeiture conditions" are a familiar, if controversial, feature of the criminal law. *See generally* Paul H. Robinson, *Causing the Conditions of One's Own Defense: A Study in the Limits of Theory in Criminal Law Doctrine*, 71 U. Va. L. Rev. 1 (1985). These conditions are easiest to understand when they're associated with, or included among the elements of, an affirmative defense. For example, a defendant whose acts and mental states would otherwise establish the elements of self-defense will nonetheless forfeit that defense if his acts and mental states also establish him to have been an "initial aggressor" or "provocateur." Statutory formulations of the provocation defense usually include, or are interpreted to include, a forfeiture rule that works in a similar manner, such that a defendant whose acts and associated mental states would otherwise establish the elements of the provocation doctrine will nonetheless forfeit the benefit of that doctrine if his acts and associated mental states also establish him to have "provoked (i.e., culpably caused) the provocation."

These familiar forfeiture conditions are similar in effect to the ones discussed in the text in connection with the partial excuse and partial forfeiture theories: they result in the loss of some otherwise available legal benefit. But they're also different. Roughly speaking, whereas these more familiar forfeiture conditions are based on something the defendant has done, together with some associated mental state or states, the forfeiture conditions discussed in the text are based on the defendant having inexcusably formed some judgment the state obligated him not to have formed (in the case of the partial excuse theory) or on being insufficiently habituated in some way in which the state obligated him to be habituated (in the case of the partial forfeiture theory).

Forfeiture rules of either the more common sort, or of the sort discussed in connection with the partial excuse and partial forfeiture theories, do not themselves "criminalize" whatever the object of the applicable forfeiture rule happens to be (i.e., an act, a judgment, an insufficiently developed habit). Instead, they deny the defendant the benefit associated with some other legal rule or doctrine when, but for the forfeiture, the defendant would have been entitled to that benefit. For example, suppose a defendant satisfies whatever it takes to qualify as an "initial aggressor" or "provocateur" for purposes of the affirmative defense of self-defense. Suppose, too, that whatever the defendant did to qualify as an initial aggressor or provocateur would not itself have amounted to a crime (though it might, depending on the definition of "aggressor" and the facts of the case). The defendant is therefore free to do whatever he did to qualify as an initial aggressor or provocateur without fear of criminal liability. But now suppose that whatever the defendant did to qualify as an initial aggressor or provocateur caused someone else to use deadly force against him. In that event, if the initial aggressor or provocateur uses deadly force to protect himself, reasonably believing that the use of such force is necessary, he'll nonetheless forfeit (to some degree) the state's permission to use deadly force, unless he'd previously "renounced" his "initial aggression" or "provocation."

18. Although the text refers for ease of exposition to the provocation doctrine as a "partial defense," sometimes the elements of the doctrine function in positive

and partial forfeiture theories are structurally the same inasmuch as they are each built upon, or constructed out of, these same two basic parts.

The excusing condition, in both theories, is a passion-induced incapacity. More specifically, the adequately provoked defendant is excused if and because passion caused him to be incapable or powerless to some extent or another to form any intent other than an intent to kill. The state, through the doctrine of provocation, recognizes this incapacity as an excuse, all else being equal. But, sometimes, all else is not equal. Sometimes, the defendant is said, in the state's judgment, to have been at fault, culpable, blameworthy, and so on for having been incapacitated at the time he formed the intent to kill.[19] If so, then, in both theories once again, the state refuses to give effect to the excusing condition. The excusing condition is "forfeited," because the defendant was at fault and so on for causing it. This forfeiture condition rests on some underlying principle to the effect that no defendant ought to benefit from a defense whose elements he's at fault for having caused or created in the first place. A more familiar example of such a forfeiture rule is the "aggressor" rule associated with self-defense, under which a defendant who would otherwise be acquitted because he acted in self-defense forfeits that defense if he qualifies as an unrenounced initial aggressor.

The partial excuse and partial forfeiture theories differ insofar as they interpret or specify these two conditions in different ways. The partial excuse theory interprets the excusing condition as partial only, and it interprets the forfeiture condition as being "all-or-nothing." When, according to the partial excuse theory, the all-or-nothing forfeiture is established, the forfeiture is "all": the partial excuse from which a defendant would otherwise have benefited is forfeited entirely. The defendant loses any and all benefit from the partial excuse and is accordingly held liable for murder. However, when the forfeiture is not established the forfeiture is "nothing": the partial excuse is not forfeited at all. The defendant thus retains the full benefit of the partial excuse. This non-forfeited partial excuse to a charge of murder yields liability for manslaughter.

The partial forfeiture theory, in contrast, interprets the excusing condition as total, and it interprets the forfeiture condition as sometimes total, and sometimes only partial. In other words, according to the partial forfeiture theory, the defendant *always* loses the full benefit, more or less, of an excuse that would have been complete or total (resulting in an acquittal) but for the forfeiture. The only question is whether the forfeiture is partial or complete. When the forfeiture is complete, the defendant loses any benefit of the full excuse and is accordingly held liable for murder. When the forfeiture is partial, however, the defendant partially loses the benefit of the full excuse and is accordingly held liable

law as a "definition" of a separate crime denominated "manslaughter." *See* App. A, cmt.1

19. Conceptual distinctions can be drawn between "being at fault," "culpability," and "blameworthiness," but those concepts are used interchangeably in the text.

for manslaughter. This partial forfeiture of an otherwise complete excuse to a charge of murder yields liability for manslaughter.

One difference between the two theories, beyond the different ways in which they interpret the defense's excusing and forfeiture conditions, is worth highlighting at the outset. Both theories portray any defendant who establishes the doctrine's elements as someone who experienced a reason-incapacitating passion; and both theories portray the state as refusing to give that incapacity mitigating effect (to some extent or another) if the alleged provocation was "inadequate." But they differ in the reasons they portray the state as giving for so refusing. According to the partial excuse theory, the state refuses to give mitigating effect to the defendant's passion-induced incapacity if (in the state's judgment) the passion the defendant experienced arose from the defendant's mistaken and inexcusable judgment or belief that he'd been seriously wronged. In contrast, according to the partial forfeiture theory, the state refuses to give mitigating effect to the defendant's passion-induced incapacity if the passion the defendant experienced arose from a perception that he'd been seriously wronged, and moreover, if (in the state's judgment) the defendant is someone without the habit of seeing or perceiving things such that he does not experience such passion.

In other words, the partial excuse theory treats a passion-incapacitated defendant as a murderer if the state identifies some "defect" in the way in which the defendant judged, or formed a belief about, the "moral" world—the world in which one person is described as having done something to "wrong" another—whereas the partial forfeiture theory treats the passion-incapacitated defendant as a murderer if the state identifies some "defect" in the way in which the defendant habitually sees or perceives that world. The partial excuse theory, one might say, scrutinizes in this way the operations of the defendant's reason; the partial forfeiture theory, in contrast, scrutinizes the habitual operations of the defendant's senses, or perhaps, his "moral sense."

The Article has four Parts. Part I gives a generic statement of the provocation doctrine. Part II describes two different psychological processes by which a defendant might cognize something as a provocation, and as a result, form an intent to kill.[20] Neither of these psychological accounts can support a theory of provocation built on the generic statement of the doctrine given in Part I. I take the time to present them

20. Existing literature on the provocation doctrine (at least the literature found in the law reviews) tends, in my opinion, to be a little thin when it comes to describing with any precision the manifold psychological processes by which a person cognizes some event in the world, experiences passion as a result, and then forms judgments and intentions, including sometimes an intent to kill, under the "influence" of passion. Indeed, I've come to believe that it makes little sense to try to articulate a theory of provocation without getting clear on the underlying psychology. In that sense, psychology precedes theory. Horder's work does more than most to probe the possible underlying psychologies. *See* Horder, *supra* note 11.

because they set the stage for Parts III and IV, which describe the partial excuse and partial forfeiture theories, respectively. The psychologies underlying these two theories, unlike the psychologies described in Part II, are capable of supporting a theory of provocation. The theories they support are, once again, similar, but also different. Parts III and IV describe these similarities and differences.

I. The Provocation Doctrine

A theory of the provocation doctrine, as understood here, is something meant to explain *why* each of the doctrine's elements is individually necessary and jointly sufficient to account for the doctrine's legal effect, which is to make a defendant liable for manslaughter when otherwise he'd be liable for murder. A theory of provocation, thus understood, aims to understand how and why the doctrine's elements, once combined, have the legal effect they have.

Any theory of the provocation doctrine necessarily presupposes, as a theory of a doctrine instantiated in positive law, some statement of its elements in positive law: some statement of that which the theory is meant to explain. The elements of the provocation doctrine differ in detail, of course, from place to place. Those differences, especially between what those elements have been taken to be at "common law," and what they are in the Model Penal Code, have been much discussed. These doctrinal differences are important. They might, depending on the facts, make the difference between a conviction for murder and one for manslaughter.

Yet, for analytical purposes — for purposes of coming up with a theory of provocation — these differences unnecessarily complicate the task. The theorist is quickly drawn into discussing differences in the scope of the defense arising from sometimes subtle linguistic differences in the language used to state the doctrine's elements in one jurisdiction compared to another. The forest can get lost for the trees. So, for present purposes, I work with what I take to be a generic statement of the provocation doctrine and its elements.[21]

The doctrine of provocation, as understood here, provides that: If (1) a person forms an intent to kill, but (a) would not have formed that intent but for having experienced a passion-induced "loss of self-control,"[22] and

21. The Model Penal Code formulation of the provocation doctrine, given in § 210.3(1)(b), has been interpreted differently in those jurisdictions that have adopted the MPC language; and some jurisdictions have, of course, adopted the MPC language with modifications. Looking only at the language § 210.3(1)(b) and the relevant Code Commentary, however, the Code's drafters understood § 210.3(1)(b) to reflect or incorporate two doctrines other jurisdictions had kept separate; namely, "provocation" and "diminished capacity." Insofar as the language of § 210.3(1)(b) extends beyond "provocation," and thus beyond the generic statement of the doctrine given in the text, the two theories discussed below will explain § 210.3(1)(b) only to the extent that § 210.3(1)(b) conforms to, but does not extend beyond, the "generic" statement.

22. The criminal law literature could, in my opinion, do a better job analyzing the

(b) would not have experienced that passion-induced loss of self-control but for having formed some cognition about some event in the world (i.e., about some "provocation"); and (2) the provocation that in fact caused the first element was, in the law's judgment, adequate; then (3) the defendant is liable for manslaughter, when absent either the first or second elements, he's liable for murder.

This generic statement may be thought controversial, especially its reference to loss of self-control. Perhaps, but the idea that a defendant must, in order to be eligible for the mitigation the doctrine provides, be found to have killed only because, in some sense, he "lost self-control" is not an uncommon feature of positive law. When explaining the Mode Penal Code's formulation of the doctrine, for example, the drafters state: "In the end, the question is whether the actor's loss of self-control can be understood in terms that arouse sympathy in the ordinary citizen."[23] Likewise, UK law makes "loss of self-control" an express element of the doctrine in both its 1957 formulation and its later 2009 formulation.[24]

"psychology of self-control." For example, the literature commonly speaks of a "capacity" for (or of) self-control. *See, e.g.*, Richard Holton & Stephen Shute, *Self-Control in the Modern Provocation Defence*, 27 Oxford J. Legal Stud. 49, 55 (2007) ("[T]here is a distinct faculty of self-control that enables agents to do what they judge best in the face of strong inclinations to the contrary.") It's fine to speak this way, but one would still like to know, for example: In what does this "capacity" consist? How is it "exercised"? What does it mean to say it can be, and sometimes is, "lost"? How does it get "lost"? Indeed, what kind of thing is a "capacity"?

According to the folk psychology on which the discussion in the text relies, see *infra* note 27, no such "distinct faculty of self-control" exists. Instead, a person "exercises self-control" through and in virtue of exercising his "powers" of "reason" and "will," which can and sometimes do work together to achieve an end or goal in a way that can be described as having exercised "self-control." In other words, a person's "capacity for self-control" is reducible to the powers of reason and will, and the "exercise of self-control" consist in nothing more than the joint exercise of these two powers, working together to achieve "self-control" as an end or goal. For more on the psychology of self-control, see *infra* pt. II.B.

Of course, the nature of a "power," "capacity," "ability," and so forth, including human powers, capacities, and abilities, is controversial, not to mention how one might come to know the extent to which, on any particular occasion, a person's power or capacity for this or that was "reduced," "impaired," diminished," "incapacitated," and so on. *See generally* Causal Powers (Jonathan D. Jacobs ed., 2017); Ruth Groff & John Greco, Powers and Capacities in Philosophy: The New Aristotelianism pt. III (2013).

23. Model Penal Code and Commentaries § 210.3 cmt.5(a) at 53 (1985).

24. Homicide Act 1957, 5 & 6 Eliz. 2 c. 11, § 3 (UK) ("[T]he question [is] whether the provocation was enough to make a reasonable man do as [the defendant] did," which was to "lose his self-control" and form an intent to kill); Coroners and Justice Act 2009, c. 25, § 54 (UK) ("Where a person ('D') kills . . . another ('V'), D is not to be convicted of murder if . . . D's acts and omissions in doing . . . the killing resulted from D's loss of self-control."). If a jurisdiction's formulation of the doctrine references only the "heat of passion," without any express reference to "loss of self-control," the phrase "heat of passion" is likely, I suspect, to be interpreted to mean, as one pair of commentors says, an "emotion of a kind and

Parts III and IV describe in more detail what the partial excuse and partial forfeiture theories require a defendant to prove to establish "loss of self-control." Appendix A offers additional reasons for the choices made in formulating the generic statement of the defense given above.

Elements (1) and (2) constitute the defense itself: those elements that, having been established, are individually necessary and jointly sufficient to establish element (3). Element (3) is not strictly speaking an element of the defense—it's a legal consequence of the elements stated in (1) and (2)—but it's nonetheless a fixed feature of existing doctrine and thus a feature any theory of the doctrine needs to explain.

Element (1) is conventionally described as a subjective element, inasmuch as it describes some psychological process by which the defendant in fact (and, in that sense, subjectively) cognized some event in the world, experienced a passion-induced "loss of self-control" as a result of that cognition; and then, as a result of having experienced that loss of self-control, formed an intent to kill. The psychological sequence reflected in element (1) thus goes from cognition to passion, to loss of self-control, and then, finally, to intent to kill. The cognized event that initiates this sequence is called a provocation.

Element (2) is conventionally described as objective, inasmuch as it describes something that must be true, in the law's judgment (and, in that sense, objectively), about the event in the world, which the defendant cognized, and which initiated the psychological process described in (1). It describes, in other words, something that must be true, in the state's judgment, about the provocation the defendant alleges. What must be true of that provocation is that it be, in the state's judgment, adequate. Because element (2) requires the state to make some judgment or evaluation about the nature of the provocation alleged to have been adequate, element (2) has also been described as a normative element. Element (1), together with element (2), then secures, by operation of law, the legal benefit described in element (3).

The two theories of provocation described below in Parts III and IV each have two parts. The first part describes a folk, commonsense, or philosophical psychology in virtue of which a defendant "loses self-control."[25] As folk psychologies, neither makes use of any specialized concepts or categories modern-day psychologists have formed and named to serve their particular purposes. Instead, it relies on ordinary mental-state concepts, like belief, desire, and intention.[26] Moreover,

to a degree that interferes with the defendant's ability to exercise self-control." Berman & Farrell, *supra* note 16, at 1043. It would take more extensive research to confirm that suspicion.

25. *See generally* Daniel Hutto & Ian Ravenscroft, *Folk Psychology as a Theory*, *in* Stanford Encyclopedia of Philosophy (Edward N. Zalta ed., 2021), https://plato.stanford.edu/entries/folkpsych-theory [https://perma.cc/4QRX-3TER].

26. *See, e.g.*, Michael Moore, Mechanical Choices 98–99 (2020):
The folk psychology in question [in the criminal law] is that relating to practical rationality. On the standard view of this psychology there are three sorts

unlike much writing in criminal law, it also relies on concepts that refer to different powers or capacities said to make up the human person. These include rational powers or capacities (specifically, reason and will), but also sensory powers or capacities, which human beings share in some way with non-rational animals.[27]

---

of representational states that cause the behavior of rational agents: there are states of desire, where we represent the world as we want it to be; states of belief, where we represent the world as we believe it is; and there are states of intention, where we represent the world as we intent to make it.

27. When the criminal law literature on the provocation doctrine looks to past thinkers for insight, it often looks to Aristotle, with reference to the *Nichomachean Ethics*. *See, e.g.*, Ashworth, *supra* note 6, at 292; R.A. Duff, *Criminal Responsibility and the Emotions: If Fear and Anger Can Exculpate, Why Not Compassion?*, 58 Inquiry 189, 202–07 (2015); Dan Kahan & Martha Nussbaum, *Two Conceptions of Emotion in the Criminal Law*, 96 Colum. L. Rev. 269, 290 (1996). Kahan and Nussbaum nonetheless report, in the just-cited article, that Aristotle's "conception" of emotion, which informs their own "evaluative conception" of emotion, was "developed also in a very influential form by Thomas Aquinas." *Id.* at 290. They make no further reference, however, to that "influential form," nor to the ways in which Aquinas "developed" it. Besides this passing statement in Kahan and Nussbaum's article, I am not aware of any reference to Aquinas's analysis of the passions, or his theory of action more generally, in connection with the criminal law literature on the provocation doctrine.

I've found Aquinas's moral psychology helpful for understanding the provocation doctrine, and that moral psychology informs the psychologies discussed in the text. Having said that, I mean for the argument presented in the text to be intelligible simply as a form of folk psychology. I limit discussion of its Thomistic roots to the footnotes. Thomist psychology is especially useful for interpreting the doctrine of provocation inasmuch as it understands human beings as constituted by distinct and complex capacities and powers. As such, it has available to it conceptual resources to present an account of human action, including "provoked" human action, that's considerably more intricate than a more conventional folk psychology based only on beliefs, desires, and intentions.

For example, a conventional belief-desire-intention psychology takes no account of what, in Thomistic psychology, are called the sensory powers or capacities of the human person. Yet once these powers or capacities are conceptually available, one can say, among other things, that "desires" arise not only from "movements" of the will (producing what Thomistic psychology would call "pseudo-passions"), but also from "movements" of the sensory appetite (producing "passions"). Likewise, a belief-desire-intention psychology takes no account of the possibility that human action results from the complex interaction among the movements of the distinct powers or capacities making up the rational and sensory parts of the human person, *see* Can Laurens Löwe, Thomas Aquinas on the Metaphysics of the Human Act 199 (2021) (For Aquinas, a "human action" is a "coordinated set of power-exercises of the agent."); and it ignores, among other such complex interactions, the interactions between "reason" and "will," *see e.g.*, John Finnis, Aquinas 71 (1998); Daniel Westberg, Right Practical Reason: Aristotle, Action, and Prudence in Aquinas 119 (1994). For an account of reason's role in the etiology of an act within Thomistic psychology, see, for example, Scott MacDonald, *Practical Reasoning and Reasons-Explanations: Aquinas's Account of Reason's Role in Action*, *in* Aquinas's Moral Theory: Essays in Honor of Norman Kretzmann (Scott MacDonald & Eleonore

The folk psychological part of each theory describes a psychological process in virtue of which a defendant, having in some way apprehended or cognized some event in the world, ends up in some sense losing self-control, and from there, forming an intent to kill. The second part of each theory is about doctrine: It explains how each theory interprets elements (1) and (2) of the provocation doctrine, given the psychological account from the first part, and then how those two elements, thus understood, combine to explain element (3).

Before describing the psychology associated with the partial excuse and partial forfeiture theories, I first describe two other psychological processes, both of which begin with the defendant cognizing some event in the world and end with his formation of a passion-induced intent to kill. Neither of these psychologies, however, can underwrite or support a theory of provocation. In both accounts, passion "disturbs" the defendant's reason, but in neither does it incapacitate his reason; and unless reason is incapacitated, to some degree at least, the defendant cannot be said to have "lost self-control," as that phrase should be understood for purposes of the provocation doctrine. I present these two accounts to distinguish them from, and to clear the way for, the partial excuse and partial forfeiture theories presented in Parts III and IV.

II. Reason Disturbed, But Not Incapacitated

The first account is here called the "hot blood" account; the second, is the "abandoned self-control" account. In both, as we'll see, passion in some way "disturbs" or "influences" the way in which the defendant's reason went about doing one of the things it does. However, neither of these psychological accounts can support a theory of provocation, because in neither account does passion produce a loss of self-control, which is an element of the doctrine that needs to be accommodated; and passion does not produce a loss of self-control because in neither account does passion "incapacitate" reason.

The hot blood and abandoned self-control accounts each describe a psychological process that begins with a defendant cognizing some

Stump eds., 1999). At a deeper level, not especially relevant here, the simple belief-desire-intention psychology, as I understand it, rests on a "metaphysics" of human acts different from that on which Thomistic psychology rests. *See* Löwe, *supra*, at 197 ("illustrat[ing] the position of [the metaphysics of] Aquinas's [action] theory relative to . . . [several] contemporary theories")

On Thomistic psychology generally, see Robert Edward Brennan, Thomistic Psychology: A Philosophic Analysis of the Nature of Man (2016 (orig. 1941)); Stephen L. Brock, Action and Conduct: Thomas Aquinas and the Theory of Action (2021); Steven J. Jensen, The Human Person: A Beginner's Thomistic Psychology (2018); Anthony Kenny, Aquinas on Mind (1993); Löwe, *supra*; Colleen McCluskey, Thomas Aquinas on Moral Wrongdoing (2017); Christopher Shields & Robert Pasnau, The Philosophy of Aquinas (2d ed. 2016); Eleonore Stump, Aquinas (2003); Norman Kretzmann, *Philosophy of Mind*, *in* The Cambridge Companion to Aquinas 128 (Norman Kretzmann & Eleonore Stump eds., 1993).

event in the world (alleged to have been adequate provocation) and ends with the formation of an intent to kill. This process is the same for each account, but only up to a point. Beyond that point, the accounts diverge. The psychological path the defendant's mind takes from cognition to intent to kill in the hot blood account differs from the path it takes in the abandoned self-control account. Each account nonetheless ends with the defendant forming an intent to kill, when he presumably would not have formed that intent had passion not disturbed his reason.

The hot blood and abandoned self-control accounts each begins with the defendant forming two cognitions. Not just any cognitions, but cognitions known as judgments. Human beings form judgments by and through the exercise of a distinctively human power or capacity. Keeping with longstanding usage, this power or capacity is known as reason. For present purposes, then, reason just is the power or capacity by and through which human beings, among other things, form judgments. First, we form judgments about (among other things) the intentions and actions of other human beings; and, second, in light of such judgments we form additional judgments about what to do, all things considered, in response.[28] All these judgments come in propositional form: they are expressed in language as propositions using more or less abstract concepts.

For purposes of understanding the psychology of provocation, the content of the first judgment is something about the nature of some event in the world. Its content can be expressed in various ways. For example, "So-and-so's β-ing treated me with extreme contempt," "So-and-so's β-ing was a serious wrong to me," "So-and-so's β-ing caused me to suffer a serious injustice," "So-and-so's β-ing was an expression of serious disrespect to me," and so on. The symbol "β" in these propositions is some description of someone's actions.

The content of the second judgment is different. It is not about the nature of some event in the world, but about what the defendant judges he should do, all things considered, about or in response to that event. The content of this second judgment is, roughly, some proposition to the effect that "I should do something to 'right' the wrong (slight, contempt, etc.) so-and-so did to me," "I should do something to make things right in response to what so-and-so did," or "So-and-so deserves something to be done to him in response to what he did to me," and so on.

For ease of exposition, I'll call both these judgments, taken together, "the judgment that θ," which might be stated in general terms along the lines of "So-and-so seriously wronged me and it would be good to achieve vindication." Note that this judgment does not yet specify *what* the defendant judges he should actually do to achieve "vindication:" It states only that he should do *something* to achieve or realize the end of

28. The former kind of judgment is typically said to result from the exercise of "theoretical" or "speculative" reason; the latter from the exercise of "practical" reason. Introducing the distinction between theoretical and practical reason into the discussion here does not seem especially important.

vindication, where vindication is something the defendant believes to be a good end and one worth taking the effort to achieve or realize.

The defendant's judgment that θ might of course be judged or evaluated from a point of view other than that of the defendant. One point of view from which the defendant's judgment that θ might be judged is, of course, that of the state. The defendant's judgment that θ might itself be judged by saying such things as, for example, it was correct or incorrect, true or false, appropriate or inappropriate, proper or improper, reasonable or unreasonable, and so on. We'll get to the language in which that judgment about the defendant's judgment is expressed for purposes of the provocation doctrine in due course. For now, the goal is to explain *what* happened in the defendant's mind, according to the hot blood and abandoned self-control accounts, not to pass any judgment on what happened in it. Judging will come later.

Anyway, having formed the judgment that θ, the defendant then, as a result, forms an intent to seek, achieve, or realize the vindication reason has determined to be the end that should, all things considered, be sought, achieved, or realized.[29] This intention in turn moves reason anew. It moves reason to begin deliberating about the means by which to achieve vindication. At the same time, however, this intention somehow causes, activates, arouses and so on a passion.[30] I use the word "pas-

29. Criminal lawyers use the word "intent" or "purpose" to refer to what Aquinas would describe as a "movement" of the will with respect to both ends and means. Aquinas uses two separate words to refer to each of these two movements. "Intentio" refers to a movement of the will with respect to ends, and "electio" with respect to means. All the various movements of the will can be generically described as "desires" (or "pro-attitudes" or "conative" attitudes, and so on), but these "desires" are not passions. They're "pseudo-passions," inasmuch as they are not "connected" to the body. *See, e.g.*, Robert Miner, Thomas Aquinas on the Passions 35 (2009). Moreover, insofar as the will is a rational appetitive power, it does not "move" except in response to some judgment of reason, i.e., except in response to some "movement" of a person's reason (his rational apprehensive power). *See, e.g.*, McCluskey, *supra* note 27, at 27 ("[A]ll activities of the will are preceded by activities of the intellect [reason]."). The text assumes the truth of this thesis, i.e., that the activities of the will are preceded by the activities of the intellect, but it's in fact contested among Thomistic scholars. *See, e.g.*, Löwe, *supra* note 27, at 92–97 (describing division between Thomistic scholars known and "voluntarists" and those known as "intellectualists" and "favor[ing] an intellectualist reading of Aquinas").

30. Aquinas would characterize a passion arising in this manner as a "consequent" passion: a passion arising after, and as a result of, some operation of reason. In the account given in the text, passion arises after reason forms the judgment that θ, but before it settles on the means by which to achieve vindication. *See, e.g.*, Daniel D. De Haan, *Moral Perception and the Function of the* Vis Cogitativa *in Thomas Aquinas's Doctrine of Antecedent and Consequent Passions*, 25 Documenti e studi sulla Tradizione filosofica medievale 289, 296 (2014) ("[C]onsequent passions can also result from . . . judgments of reason" prior to the "final judgment"); Steven J. Jensen, *Virtuous Deliberation and the Passions*, 77 The Thomist 193, 227 (2013) (noting that consequent passions can influence reason's judgments at different points in a deliberative process).

sion" to describe that which the defendant's intent to attain vindication arouses, rather than the more familiar word "emotion," not to emphasize or highlight the thing's force, intensity, vehemence, and so on, but rather to emphasize its passivity. Passions do not arise on their own. They arise only in response to some cognition.[31] This originating cognition in both the hot blood and abandoned self-control accounts is the judgment that θ.[32]

Aquinas characterizes the passion described in the text as arising from the "overflow" of the will into the sensory appetite. Summa Theologica, I-II.24.3 ad.1. Aquinas is "vague about the precise mechanism of this overflow." Nicholas E. Lombardo, The Logic of Desire: Aquinas on Emotion 90 (2011). The problem is how passion can arise from the formation of an intent if one assumes that passion (i.e., movements of the sensory appetite) arises only in response to some movement of the sensory apprehension. *See, e.g.*, De Haan, *supra*, at 317 (The sensory appetite is "always specified by the cognitive objects apprehended" by the sensory apprehension.); Matthew James Dugandzic, A Thomistic Account of the Habituation of the Passions 139 (2019) (unpublished Ph.D. dissertation, Catholic University of America) (on file with author) ("The cogitative power [which is part of the sensory apprehension] is the sole proximate mover of the sensitive appetite for all of its movements, except pleasure and pain."). One possibility is that the "will first moves the intellect [reason] by the vehemence of its affections [a "desire" or "pseudo-passion"] regarding some object [e.g., vindication], so that the intellect causes the particular reason [i.e., one power composing the sensory apprehension] to form an intentional object that engages the passions [e.g., causes anger]." Lombardo, *supra*, at *id. But see* De Haan, *supra*, at 297 (suggesting an alternative mechanism).

31. The text adopts the thesis that "passions are not themselves cognitive states, they are responses to cognitive states." Claudia Eisen Murphy, *Aquinas on Our Responsibility for Our Emotions*, 8 Medieval Phil. & Theology 163, 167 (1999); *see also* Christopher A. Bobier, *Thomas Aquinas on the Relation Between Cognition and Emotion*, 86 The Thomist 219, 242 (2022) ("Cognitive acts are essential to bring about, direct, and sustain emotions but they are not constituent parts of emotions proper.").

32. Aquinas draws a distinction between "consequent" passions and "antecedent" passions. Consequent passions arise consequent to (and therefore after) the operation of reason (e.g., in the formation of some judgment); antecedent passions arise antecedent to (and therefore before) the operation of reason. *See, e.g.*, McCluskey, *supra* note 27, at 102; De Haan, *supra* note 30, at 294 n.14. The distinction between antecedent passion and consequent passion is centrally important to the different psychologies on which the partial excuse theory and the partial forfeiture theory rest. The psychology on which the partial excuse theory rests presupposes a consequent passion (as do the "hot-blood" and "abandoned self-control" accounts); the psychology on which the partial forfeiture theory rests presupposes an antecedent passion.
For general discussions on how "passions" figure into Thomistic psychology, see, for example, Dina Fritz Cates, Aquinas on the Emotions (2009); Nicholas Kamm, Aquinas on Emotion's Participation in Reason (2019); Miner, *supra* note 29; Lombardo, *supra* note 30; Peter King, *Emotions*, *in* Oxford Handbook of Aquinas (Brian Davies & Eleonore Stump eds., 2012); Henrik Lagerlund, *The Systemization of the Passions in the Thirteenth Century*, *in* Philosophy of Mind in the Early and Hight Middle Ages 157, 164–74 (Margaret Cameron ed., 2018) (Aquinas "gives by far the most comprehensive treatment of the passions in the whole medieval tradition."); Dugandzic, *supra* note 30.

The passion aroused in this way can itself be called by different names. The English language includes a large number of words capturing different species of this generic passion, e.g., irritation, annoyance, pique, anger, resentment, indignation, outrage, fury, wrath, rage, and so on. The variety is striking. For now, I'll call the relevant passion, simply and generically, "anger,"[33] without meaning to imply anything with regard to its "force," where a passion's force just is the extent to which changes or alters the way a defendant's reason would have worked but for the passion. Thus, a defendant who forms the judgment that θ, and who experiences anger as a result, is someone who's in fact been "provoked," where being "provoked" means experiencing the passion of anger. But someone who's in fact been provoked is not yet someone who's "lost self-control." Being provoked is not, at least not for present purposes, the same thing as losing self-control.

33. Two ancillary comments on the Thomistic account of "anger" and how it can be seen to relate to the doctrine of provocation. First, according to Thomistic psychology, "anger" results from the "cognition of an arduous present evil [malum] with the hope of overcoming it." *See, e.g.*, LOMBARDO, *supra* note 30, at 63. In a case involving an alleged provocation, the "arduous present evil" is the wrong already done to the person, the cognition of which produces the passion of "sorrow"; the way in which this wrong is "overcome," and thus that which is the object of hope, is vindication, cognized as a good. If vindication is achieved, the passion of anger comes to rest in "pleasure." *See, e.g.*, MINER, *supra* note 29, at 86. In short, according to Thomistic psychology, anger (an irascible passion) begins in sorrow (a concupiscible passion) and ends in pleasure (another concupiscible passion).

Second, in some jurisdictions, the law permits a jury to return a manslaughter verdict if it finds that the defendant experienced some emotion other than "anger." The most common alternative appears to be "fear." *See, e.g.*, JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW § 31.07[B][1], at 517 (9th ed. 2022) (citing cases). In jurisdictions not recognizing fear as an alternative to anger, some commentors urge those jurisdictions to change their law to make it an alternative. *See, e.g.*, Michal Buchhandler-Raphael, *Fear-Based Provocation*, 67 AM. U. L. REV. 1719 (2018).

As a matter of Thomist psychology, however, it would appear to be a conceptual mistake to include fear as a "predicate" passion upon which to base a plea of provocation. Thomistic psychology supposes that "fear" arises upon some cognition of a "*future* evil impossible to overcome." Yet inasmuch as the provocation doctrine requires of any alleged provocation, whether judged "adequate" or not, that it be something *already* done to the defendant, fear ca not coherently be the passion on which the doctrine rests or should rest. If so, then in cases in which a provoked defendant reports having experienced fear, or in which fear is ascribed to the provoked defendant, the defendant should nonetheless be required to establish having cognized *something* amounting to an "arduous present evil." The defendant may indeed have experienced fear or some other passion, along with anger. The experience of one passion need not exclude the experience of others. Indeed, the experience of one passion can cause the experience of other passions. Nonetheless, anger is the passion that moves toward vindication; fear is the passion that moves away from the thing feared. As such, the "logic" of fear does not "fit" the vindicatory action associated with the doctrine of provocation.

Nor is someone who's in fact been provoked yet someone who's formed a further judgment making killing the all-things-considered means by which to achieve vindication as the end. That judgment remains to be made. Of course, because we're dealing with the provocation doctrine, we're dealing with a psychology in which the relevant sequence of events does indeed terminate in the defendant forming an intent to kill. The question is what happens between the passion of anger getting aroused and the formation of an intent to kill. It's at this point, after or upon the arousal of passion, that the hot-blood account and the abandoned self-control accounts diverge. They diverge inasmuch as each offers a different explanation for how the defendant goes from experiencing anger to forming an intent to kill.

A. *Intending to Kill in Hot-Blood*

According to the hot-blood account, the defendant's will, in virtue of having formed an intent to seek vindication as an end (itself resulting from the judgment that θ), causes the defendant's reason to deliberate about, or balance, reasons for or against various means by which to achieve that end. At the same time, the formation of an intent to seek vindication causes the defendant to experience the passion of anger. In other words, while reason is trying to sort out what to do to achieve vindication, passion is at work influencing or disturbing reason's deliberations, causing it to balance reasons for and against this or that means in a way it would not have gone about that balancing had passion not been aroused.

For any defendant claiming provocation as a defense, the anger he experienced was presumably, in some sense, strong, powerful, intense, vehement, etc. What those words mean for a hot-blooded defendant is that passion somehow causes reason, as it goes about deliberating over means, to give greater weight to reasons to kill, and less weight to reasons not to kill. Reasons to kill might be said, for example, to become somehow more salient as a result of passion than do reasons not to kill, causing reason to attend to, or focus on, reasons to kill, and giving less attention to reasons not to kill. As a result of passion, reasons to kill somehow carry more weight in the defendant's practical deliberations than they otherwise would have.

In this way, the hot-blooded defendant ends up forming an all-things-considered judgment making killing the means to vindication, when, without such a passion-produced disturbance of reason, reason would not have made killing the means to vindication. Then, in the final movement, with reason having settled on killing as the means that should be chosen to secure vindication, the defendant's will, consistent with reason's judgment, chooses to kill; or, in language more familiar to the criminal law, the defendant forms an intent to kill.

The psychological process thus described, beginning with the judgment that θ and ending with an intent to kill, has said nothing about the defendant losing or having "lost self-control." The provocation doctrine's

first element, however, requires the defendant in some sense to have lost self-control, such that he ended up forming an intent to kill. Consequently, because the psychology of a hot-blooded defendant includes nothing amounting to a loss of self-control, the psychology of a hot-blooded defendant is insufficient to establish a necessary element of the defense. Such a defendant kills in hot blood, and might therefore not be liable for first-degree murder, but not having in any sense lost self-control, a verdict of manslaughter is off the table.

The hot-blooded defendant, one might say, went "overboard," "over-reacted," or "acted disproportionately," and so on, when he chose to kill in response to the alleged provocation. Saying a defendant went overboard, over-reacted, or acted disproportionately presupposes some standard against which to measure or assess what he chose to do in response to the alleged provocation, where what he chose to do was to kill. Such a standard implicitly assumes that *some* responses to the alleged provocation would have been a proportionate, and thus presumably permissible or appropriate means to achieve the end of vindication for someone who'd experienced the provocation alleged. Any response beyond those permitted responses would then be described as disproportionate. The state, of course, reserves for itself the judgment as to how serious a wrong the alleged provocation was; what response, if any, it would have permitted as a means to achieve vindication; and what responses it will not permit.[34]

B. *"Abandoning Self-Control" and Intending to Kill*

A defendant who kills in hot-blood is someone who, as a result of passion, forms an all-things-considered judgment that killing is the means to be chosen to achieve vindication as an end, which judgment in turn causes him to form an intent to kill. A defendant who abandons self-control likewise forms an intent to kill, but the psychological path to

34. The work of some contemporary commentators can, I think, fairly be read to suppose that the psychology of hot blood is indeed the psychology on which the provocation doctrine rests, but I would not try to make the case for that reading here. *See, e.g.*, Kahan & Nussbaum, *supra* note 27; Nourse, *supra* note 1; *see also* Horder, *supra* note 11, at 59–71 (arguing that the early common law was based on a psychology the author labels "anger as outrage"). A theory of provocation based on this psychology would interpret element (2) of the provocation doctrine as a doctrinal category by which the state sorts defendants into, on the one hand, those whose choice to kill was, as Horder puts it, disproportionate to the alleged provocation, but not "greatly" so (in which case the alleged provocation would be called "adequate provocation"); and, on the other, those whose choice to kill was not only disproportionate to the alleged provocation, but "greatly" so (in which case the alleged provocation would be called "inadequate provocation"). If the psychology of hot blood was the psychology on which early common law judges built the provocation doctrine (which is how I read Horder's historical analysis), then commentators like Kahan, Nussbaum and Nourse can perhaps be understood to be urging a return to the common law, provided of course that the norms used for judging the adequacy of an alleged provocation are the ones they endorse, and not whatever norms the common law judges endorsed.

the formation of that intention, once passion has been aroused as a result of the judgment that θ and the intent to seek vindication, is different and more complicated. Still, like the hot-blooded defendant, the defendant who abandons self-control is not eligible for the mitigation associated with the provocation doctrine. A defendant who abandons self-control has not "lost" the power or capacity by which a person "controls" the causal influence a passion otherwise would have had on the intentions he forms absent the exercise of that capacity. Instead, he's chosen or decided to stop exercising that capacity, and in that sense, he has "abandoned self-control."

Like the hot-blooded defendant, the defendant who abandons self-control experiences a reason-disturbing passion. Unlike the hot-blooded defendant, however, the defendant who abandons self-control does not at the outset judge that killing should be the means chosen to achieve vindication. Indeed, despite the effects of passion on reason, he forms, at the outset, the opposite judgment: he judges all things considered that killing should *not* be the means chosen to achieve vindication.

Instead, he judges that he should do something else, besides killing, to secure vindication, if he judges he should do anything at all. Maybe he judges that he should yell or shout or hurl insults at the person he judged has wronged him. Maybe something along those lines, he judges, is the best way to achieve vindication. Or maybe he judges that he should threaten him in some way, but nothing more than that. Or maybe, even, he judges that he should cause him some physical harm, short of killing him. But whatever the means he judges should be the means he should choose all things considered to achieve vindication, or even if he forms no judgment one way or the other about means, he also judges that he should not choose to kill for the sake of vindication.

Alas, despite this judgment, itself formed despite the influence of passion, passion has not yet subsided. It persists, pushing reason to reconsider and rebalance the reasons he has for and against killing. Passion wants reason, so to speak, to change its mind, to shift from its judgment against killing to a judgment in favor of killing. Realizing passion might cause such a shift, such that killing would become the means chosen to achieve vindication, the defendant's reason forms yet another all-things-considered judgment: In order not to end up forming an intent to kill, reason judges that something should be *done* to control or contain this persistent passion, lest it cause him to change his mind in favor of killing as the means to vindication. Controlling might here be understood to mean calming: doing something to diminish the causal effect passion would otherwise have on reason.

But what should the defendant do, now that controlling passion has become his end? What means does he judge he should pursue to achieve the end of controlling passion, which end is itself a means to achieve his more remote end of not killing?

The available options include any number of familiar acts: counting to ten, taking a deep breath, bringing happy thoughts to mind, bringing to mind thoughts of the consequences if passion "gets the upper hand," just walking away, and so on. Doing all these acts—many of which are mental acts—is what it means, in one sense of the phrase, to exercise self-control, where the power being exercised is not any special, distinct, or sui generis human power. Exercising self-control involves nothing other than exercising the powers of reason and will, working together to produce actions (including mental ones) the intended effect of which is to calm the passions.[35] Self-control is exercised when reason and will work together, in the ways they do, first to contain the causal influence passion would otherwise exercise on reason, and then somehow to weaken or diminish its causal influence. If the defendant performs these sundry acts of self-control, he might succeed in achieving self-control, in which case he'll not change his mind about not killing, in which case he would not form an intent to kill.

But he might also fail. Exercising self-control, especially in the face of a powerful passion, is typically experienced as an effort or a struggle, as something that's hard or difficult, producing "stress" or "dysphoria."[36] These unwelcome effects of the struggle between reason and passion then become a reason not to continue it. At some point, the balance of reasons the defendant's reason receives and to which it responds, for and against continuing to do all the things in which the exercise of self-control consists, might shift. Having judged at the start that he should exercise self-control, thereby at least containing passion, the defendant's

35. A Thomistic account of the "mental acts" referred to in the text might be as follows. Reason moves the sensory apprehension to form "perceptions" (broadly understood to includes memories and imaginings), which in turn cause "countervailing" passions that somehow reduce or diminish the influence anger would otherwise have had on reason. In other words, a person's reason judges he should "bring to mind" (i.e., move the sensory apprehension to produce) certain "perceptions" to diminish passion's influence on reason; the will, in response to this judgment, causes the sensory apprehension to produce those "perceptions"; those "perceptions" then in turn move the sensory appetite to produce countervailing passions. *Cf.* Löwe, *supra* note 27, at 178–92 (discussing the sensory power of memory). This way of "exercising self-control," according to a Thomistic psychology, does not seem to differ all that much from standard analyses of "exercising self-control" found in the non-Thomistic philosophical literature; for example, Alfred Mele, Irrationality: An Essay on Akrasia, Self-Deception and Self-Control 26 (1987) (discussing "skilled" self-control), minus any reference to the various powers or capacities integral to Thomistic psychology. *See also* Andrea Scarantino, *Exploring the Roles of Emotions in Self-Control*, *in* Surrounding Self-Control 116, 138 (Alfred Mele ed., 2022) (noting that "affective strategies" can "harness the distinctive power of emotions to work for self-control rather than against it").

36. Insofar as this experience, described as "stress," "dysphoria," and so on, has physical or bodily manifestations or consequences, one suspects it must have something to do the movements of the sensory appetite. That's because, in Thomistic psychology, the sensory powers (including the sensory appetite) are, unlike the rational powers of reason and will, somehow "connected to" the body.

reason might revise that judgment as the struggle against passion continues. Reason might end up reversing course, moving from the judgment to make exercising self-control the means to achieving the end of not killing, to the judgment to stop exercising self-control as a means to avoid the unpleasantness attendant to its continued exercise.[37]

If the defendant's reason does indeed judge that the exercise of self-control should be abandoned, he (his will) will no longer continue to form (or sustain) the intent to contain or diminish passion (i.e., to exercise self-control); and without that intent, he will stop doing all the things he had been doing to keep passion in check. Whatever the defendant had been doing to prevent passion from influencing reason's judgment not to kill will come to an end. He stops resisting passion, and with that resistance abandoned, passion regains the power to move reason as reason deliberates about the means to be chosen to achieve vindication. Reason, now under passion's influence, will move from its all-things-considered initial judgment—"I should not kill"—to a revised all-things-considered judgment—"I should kill." The defendant (his will) will then, in accordance with what reason instructs, form an intent to kill (albeit in hot blood).

When a defendant abandons self-control, he starts with the judgment that he should not kill, but nonetheless eventually "changes his mind" in favor of its opposite: that he should kill. As such, a defendant who abandons self-control might be variously described as having acted with a "weak will," "akratically," or "incontinently." Describing the defendant's choice to kill in those ways is harmless enough, but caution remains in order. Words like "weak-willed," "akratic," and "incontinent," standing alone and without elaboration, may obscure the psychological process by which a defendant chooses to stop doing whatever he was doing to keep passion in line. Considerable controversy and confusion surround the nature and intelligibility of the psychological state words like "weak-willed," "akratic," and "incontinent" are meant to name, as well as acts performed while in that state.[38] What matters is the psychological process itself, not the label used to name it.

37. A criminal lawyer might at this point ask if the defendant's judgment that he should stop "exercising" self-control was "reasonable." Although the positive law in some places may be otherwise, the generic doctrine of provocation stated in the text does not require a finding that the defendant's "loss of self-control" was "reasonable." For more on this, see *infra* app. A, cmt. 9.

38. A defendant who forms an intent to kill because he "abandons self-control" in the manner described in the text is not "weak-willed," "akratic," or "incontinent" in at least one sense in which those words are commonly used. A "weak-willed" choice is sometimes described as a choice made against, or contrary to, a person's all things considered judgment about what he should do. In other words, a choice is "weak-willed" just in case the person judges all things considered that he should not θ, but nonetheless forms an intent to θ. But that's not true of a defendant who "abandons self-control" in the way described in the text. A defendant who "abandons self-control" in the way described in the text has, as a result of its abandonment, formed the all-things-considered judgment that he

Still, however labeled, a defendant who abandons self-control in the manner described above might be thought less culpable for having formed an intent to kill, at least when compared to a hot-blooded defendant, who made no choice to exercise self-control at all. A defendant who abandons self-control at least made some effort, albeit unsuccessful, to contain or diminish the causal influence of passion on reason. Nonetheless, like the hot-blooded defendant, the defendant who abandons self-control will, under existing law, be liable for murder. Again, the doctrine of provocation's first element requires the defendant to have "lost self-control." A defendant who abandons self-control has not "lost" it: he's abandoned it, or at least he's abandoned its continued exercise.

Now would be a good time to summarize. A defendant who, as a result of passion, forms an intent to kill in hot blood has not lost self-control. Nor has a defendant who's abandoned self-control. Neither has lost self-control because, despite the experience of passion, neither has "lost" a capacity or power necessary for exercising self-control;[39] namely, the power of reason, and the defendant's power of reason has not been lost because it has not to any degree been incapacitated, disabled, or disempowered. Yes, passion has disturbed the power of reason, causing it to form a judgment about means it otherwise would not have formed. And, yes, but for passion's effect on reason, an intent to kill would not have been formed. But, on the account of how passion incapacitates, disables, or disempowers reason offered below,[40] passion simply has not done any of those things to the reason of a defendant who forms an intent to kill in hot blood, or after having abandoned self-control.[41]

---

should kill, and then, as a result of that judgment, forms an intent to kill. His intention is therefore not "out of line" with his judgment.

According to one commentator, Aquinas would characterize an intent formed "from passion or weakness" as one formed as a result of passion having "incline[d] the intellect [reason] against its own knowledge, for the particular judgment [to θ] is itself contrary to what is known in the universal." Bonnie Kent, Virtues of the Will: The Transformation of Ethics in the Late Thirteenth Century 159 (1995). In other words, a "weak-willed" act is one in which the agent forms the intent to θ because he judges "in the particular" that he should θ, while at the same time judging "in the universal" that he should not θ. *See, e.g.*, McCluskey, *supra* note 27, at 113 ("Aquinas characterizes the situation of the incontinent agent as a failure of knowledge of particulars."). This account preserves a sense in which a "weak-willed" actor intends to θ while at the same time judging ("in the universal") that he should not θ.

39. Reason is necessary for the "exercise of self-control," but not sufficient. "Exercising self-control" also requires the will to do what reason instructs it to do, i.e., among other things, to move the sensory apprehension. *See supra* note 35 (describing how "self-control" gets "exercised").

40. *See infra* pt. III.A.

41. As stated in the text, the "hot blood" and "abandoned self-control" psychologies ca not support a "theory" of provocation because they ca not explain its "loss of self-control" element. Still, they might nonetheless be thought to give anyone, including the state, reasons for believing that a defendant who forms an intent to kill in the way each account describes is somehow "less culpable" compared to a

III. The Partial Excuse Theory: Reason "Partially Incapacitated"

Unlike the hot-blood and abandoned self-control accounts, the partial excuse and partial forfeiture theories of provocation rest on, or presuppose, a psychology in which the defendant does indeed, to some degree, lose self-control. Passion not only disturbs reason, it also incapacitates, disables, or disempowers it. With reason being more or less incapacitated, disabled, and so on, the mental acts needed to exercise, let alone achieve, self-control would not happen because, to some degree, they ca not. Passion has, to one degree or another, seen to that. In other words, the way passion works in the psychologies on which the partial excuse and partial forfeiture theories rest is especially devious: passion disables or undermines the very power – the power of reason – needed to control it.[42]

Although the partial excuse and partial forfeiture theories both rest on a psychology in which passion incapacitates reason, they differ in other ways, two of which are perhaps worth previewing at this point. First, in the partial excuse theory, the cognition producing a reason-incapacitating passion is a *judgment*. The defendant forms the judgment that seeking vindication would be a good end to pursue because the decedent seriously wronged him. In the partial forfeiture theory, in contrast, the cognition producing a reason-incapacitating passion is *something else*. What this something else is will be described in due course in Part IV. Second, in the partial excuse theory, passion is assumed to have only *partially* incapacitated the defendant's reason; in the partial forfeiture theory, in contrast, passion is required to have *totally* incapacitated it.

A. *Psychology*

The psychological process ending in the defendant forming an intent to kill is, according to the partial excuse theory, much like that of both the hot-blood and the abandoned self-control accounts, but it differs from them in an important way. According to the partial excuse theory, the defendant, through the exercise of reason, forms the judgment that θ: that "So-and-so seriously wronged me and it would be good to achieve vindication." This judgment in turn produces an intent to seek vindication, which in turn arouses or moves passion.[43] So far, so good: the

---

defendant who forms an intent to kill absent passion having had any comparable influence on reason's operation. Indeed, one way to understand the mitigation historically associated with what Horder labels "anger as outrage," *see* Horder, *supra* note 11, at 59–71, and my own prior, but mistaken, effort to ground the provocation doctrine on a psychological account of "weakness of will," *see* Garvey, *supra* note 16, at 1729, would be to see them, not as theories of the doctrine of provocation (as that doctrine is stated in the text) but as possible alternative grounds for some non-provocation mitigation.

42. *Cf.* Holton & Shute, *supra* note 22, at 58 (describing how passion can sometimes "undermine" self-control).

43. *See infra* pt. III.A.

psychology on which partial excuse theory rests is as yet no different from the psychologies associated with hot-blood and abandoned self-control.

But once passion is aroused in this way, its effect on reason, in the case of the partially excused defendant, is different and more profound. For a defendant who forms an intent to kill in hot blood, or who forms an intent to kill after abandoning self-control, passion "disturbs" reason, in the sense of causing it to pay more attention, so to speak, to reasons to kill, and less attention to reasons not to kill. In contrast, the partial excuse theory supposes that passion's effect on reason is more pronounced. Passion does not simply distort the weight or salience of the reasons to which reason attends, or which it receives. Instead, passion limits or restricts, to some extent, the range of reasons available to reason. Rather than simply altering the weight or salience of the reasons to which reason attends, it prevents reason from attending to, or receiving, some range of reasons not to kill altogether.[44] Being in this incapacitated state, reason then forms the judgment making killing the means to securing vindication.

This reasons-limiting effect passion has on the reasons available to reason is what the partial excuse theory means when it says a defendant's reason was, as a result of passion, more or less incapacitated, disabled, disempowered, and so forth. Any such incapacity is a matter of degree. The more reasons against making killing the means to vindication that passion manages to prevent reason from attending to, and the weightier those excluded reasons, the more passion can be said to have incapacitated reason. Any reason to which reason has not attended is, of course, necessarily a reason to which reason does not respond. Finally, because reason is a capacity or power necessary for exercising self-control, the same passion that disables reason's power to deliberate about means also disables its power to exercise self-control.

One might describe the occasions on which passion has such an incapacitating effect on reason using more colloquial phrases. The

44. A defendant "loses self-control" just in case passion "incapacitates" reason in the manner described in the text. This "loss of self-control" might be described in various other ways: "impaired capacity for self-control," "impaired capacity for reason," "impaired capacity for rationality," "diminished capacity for rationality," "impaired rationality," "impaired judgment," and so on. What's important here is not the label, but the underlying psychology in virtue of which passion is said to "incapacitate" reason. Of course, a person's "capacity" for "reason" or "rationality" can be "impaired" or "diminished" for reasons having nothing to do with passion, as with the "diminishment" or "impairment" associated with the doctrine of "diminished capacity," for example.
I am not entirely confident that the description given in the text as to how passion "incapacitates" reason is accurate or correct. I mean for that description to be substantially the same as that found in Moore, *supra* note 26, at 337–40, and Stephen J. Morse, *Moore on the Mind*, *in* Legal, Moral and Metaphysical Truths: The Philosophy of Michael S. Moore 233, 244–45 (Kimberly Kesler Ferzan & Stephen J. Morse eds., 2016), but I've come to suspect that more needs to be said to make sense of the idea that, as Moore puts it, certain desires (or passions) are such that they "refuse to be integrated into one's sense of self." Moore, *supra*, at 337. But saying more will need to wait for another occasion.

defendant, one might say, "could not think straight," or "could not think of anything else except killing." Or, in more metaphorical phrases historically associated with the provocation doctrine, the defendant's reason was "blinded by passion;" passion rendered the defendant "deaf to the voice of reason;" passion "drowned out the voice of reason;" or passion was such that the defendant was not the "master of his mind," and so on. Yet however passion's incapacitating effect on reason is described, the result is that reason forms a judgment—specifically, to make killing a means to vindication—it otherwise would not have formed, thanks to passion having altogether excluded from reason's domain some range of reasons not to kill.[45]

B. *Doctrine*

If a defendant lost self-control in the way the partial excuse theory says he lost it, what would that mean for the provocation doctrine, and in particular, what would it mean for the way each of the doctrine's elements should be understood or interpreted?

Here, by way of reminder, are the doctrine's three elements, which any theory of provocation needs to interpret and fit together such that proof of elements (1) and (2) explain element (3). The doctrine of provocation provides that: if (1) a person forms an intent to kill, but (a) would not have formed that intent but for having experienced a passion-induced loss of self-control, and (b) would not have experienced that passion-induced loss of self-control but for having formed some cognition about some event in the world (i.e., about some provocation); and (2) the provocation that in fact caused the first element was, in the law's judgment, adequate; then (3) the defendant is liable for manslaughter, when, absent either the first or second elements, he is liable for murder.

1. Element (1) – Loss of Self-Control

According to the partial excuse theory, a defendant loses self-control, for purposes of element (1), when passion incapacitates reason in the manner specified above. When a defendant forms an intent to kill but could not, due to a passion-induced incapacity of reason, have intended otherwise, the provocation doctrine excuses him for having formed that intent, *provided* element (2) is established. Element (1) thus functions as an excusing condition. The extent to which the partial excuse theory supposes a defendant's reason to have been incapacitated, and thus the extent to which it supposes he lacked the capacity not to form an intent to kill, would not be specified until we get to element (3).

45. Stephen Morse has long been making the case that "strong" emotions or passions exculpate, when they do, as a result of their effect on reason. *See, e.g.*, Stephen J. Morse, *Diminished Rationality, Diminished Responsibility*, 1 Ohio St. J. Crim. L. 289, 295 (2003) ("[I]n virtually all cases in which a defendant presents a plausible claim for a pure control excuse [i.e., an excuse based on the experience of a "strong" emotion or passion], careful analysis demonstrates that the claim collapses into an irrationality claim.").

2. Element (2) – Adequate/Inadequate Provocation

Reason-incapacitating passions do not arise from nothing. No passion does. For purposes of the partial excuse theory, passion results from the formation of the judgment that θ, together with the intent to seek vindication.[46] Pursuant to the provocation doctrine's second element (the adequate provocation element), that which causes the defendant to form the judgment that θ is that which the state judges to have been "adequate" provocation or not, as the case may be. How does the partial excuse theory interpret this element?

First, element (2), according to the partial excuse theory, is a doctrinal category by and through which the state judges the defendant for having formed the judgment that θ: the judgment that the decedent seriously wronged the defendant and that the defendant should do something to achieve vindication. In other words, the defendant's judgment that θ is itself subject to the state's judgment, in virtue of element (2).[47] The defendant's judgment that θ is the object of the state's judgment because, but for having formed that judgment, the defendant would not have experienced the reason-incapacitating passion he in fact experienced; and, of course, but for having experienced that reason-incapacitating passion the defendant would not have formed an intent to kill.

Second, the partial excuse theory's interpretation of element (2) presupposes that anyone subject to the state's jurisdiction is obligated not to form any judgment causing him to experience a reason-incapacitating passion. No matter how serious the defendant judged the wrong to which he'd been subject, and no matter how great a good the defendant judged vindication to be, those judgments are not judgments the state permits anyone to form if, as a result, they produce a reason-incapacitating passion. Of course, any defendant who interposes a plea of provocation has necessarily breached that obligation: he's formed the judgment that θ, and as a result, experienced a passion of such force as to incapacitate his reason; otherwise, he would not have established element (1).

Third, and here's the main point, the state will, pursuant to element (2), nonetheless excuse some defendants for having formed the judgment that θ, despite its having produced a reason-incapacitating passion. If the state decides that the provocation alleged to have caused that judgment is adequate, then the defendant will be excused for having formed it. If the state excuses the defendant for having formed the judgment that θ, then so far as the state is concerned the defendant was not "at fault," "culpable," "blameworthy," and so on for experiencing the reason-incapacitation passion he in fact experienced. Conversely, if the state decides that the provocation alleged to have caused that judgment

46. *See supra* note 30 (discussing "overflow").

47. How this judgment gets made, and the identity of the various state actors involved in making it, will of course depend on the specifics of each jurisdiction's provocation doctrine and the procedural rules for adjudicating claims based on it.

is inadequate, then the defendant will not be excused for having formed it. If the state refuses to excuse the defendant for having formed the judgment that θ, then so far as the state is concerned the defendant was "at fault," "culpable," "blameworthy," and so on for the experiencing the reason-incapacitating passion he in fact experienced.

Fourth, if an alleged provocation is found to have been "inadequate" and the defendant is thus found to have been at fault and so on for having formed the judgment causing him to experience a reason-incapacitating passion, then the state will deny, or render forfeit, the excuse it would otherwise have recognized in virtue of the defendant having established element (1). Conversely, inasmuch as an alleged provocation is found to have been "adequate" and the defendant is said not to be at fault and so on for forming the judgment causing him to experience a reason-incapacitating passion, the state will not deny, or render forfeit, the excuse the defendant has established in virtue of having established element (1). In other words, element (2) functions as a forfeiture condition, which is established when the provocation is found to have been "inadequate," but not when the provocation is found to have been "adequate."

Legal rules or tests to identify, with more or less precision, when an alleged provocation is "adequate" are many and varied. For example, an alleged provocation has, to take a well-known test, been said to be adequate if it would have caused a "reasonable" person to lose self-control.[48] Sometimes the phrase "reasonable person" is replaced with "ordinary person."[49] Likewise, sometimes the word "would" is replaced with "might."[50] Under these familiar rules, if a defendant forms the judgment that θ, and experiences a reason-incapacitating passion as a result, then whatever caused him to experience such a passion will be judged adequate, if and only if, in the state's judgment, a reasonable or ordinary person would or might likewise have formed that judgment.

Recent academic proposals have approached the question from the opposite direction. Rather than purporting to identify what should make an alleged provocation "adequate," these proposals purport to identify what should make an alleged provocation "inadequate." For example, an alleged provocation should, it's been said, be inadequate if, in the state's judgment, it "contradict[s] the fundamental values of the political community,"[51] or if it "offend[s] the norms and values of the community within the jurisdiction,"[52] or if it's inconsistent with "the views of the people of the state."[53] Other proposals identify more specific events as

48. Homicide Act 1957, 5 & 6 Eliz. 2 c. 11, § 3 (UK).
49. *See, e.g.*, Dressler, *supra* note 2, at 998.
50. *See, e.g., id.*
51. Jonathan Witmer-Rich, *The Heat of Passion and Blameworthy Reasons to Be Angry*, 55 Am. Crim. L. Rev. 409, 414 (2018).
52. Paul H. Robinson & Lindsey Holcomb, *Individualizing Criminal Law's Justice Judgments: Shortcomings in the Doctrines of Culpability, Mitigation, and Excuse*, 67 Vill. L. Rev. 273, 322 (2022) (emphasis omitted).
53. Peter Westen, *Individualizing the Reasonable Person in Criminal Law*, 2 Crim. L.

being "inadequate" provocation as a matter of law,[54] or provide that all alleged provocations are "inadequate" unless the decedent's action itself amounted to a "legal,"[55] or "criminal,"[56] wrong.

Of course, the more abstract or general the language in which such rules or tests are formed, the more room they leave for disagreement among whoever the law tasks with applying that language to the facts of a particular case, subject to appellate review for sufficiency or failure to instruct. Anyway, the language the law should use to sort alleged provocations into the adequate or inadequate category is in the end a question for political authorities. The state gets to decide, even if it decides to leave the decision more or less to the judgment of a jury.

What makes provocation adequate or not has changed with the social and political times, and will no doubt continue to change as the times change. Sometimes such change happens case-by-case, reflecting the changing judgments of judges and juries applying existing law. Sometimes such change happens as a result of legislative action, making changes to existing law. And, for better or worse, political authorities have an abundance of academic commentators eager to offer (often conflicting) advice on how the state should change the law to keep with the times, or perhaps on how it should change the law to get ahead of them.

Setting aside whatever specific language the state should use to distinguish "adequate" provocation from "inadequate" provocation, element (1) and element (2) structurally combine to make the provocation doctrine, as the partial excuse theory portrays it, what might be called a double excuse. Element (1) can be described as a first-order excuse. When element (1) is established, the defendant is excused to some extent, subject to element (2), for having formed an intent to kill. Element (2) can also be described as an excuse, albeit a second-order excuse. It constitutes a second order excuse inasmuch as it is an excuse that regulates when the state will deny a defendant access to the first-order excuse inscribed in element (1). The two excuses have different objects but

& PHIL. 137, 159 (2008). Westen adds that a state might elect to tell a jury, with some greater degree of specificity, just what are "the views of the people of the state." *Id.*

54. *Cf.* N.Y. PENAL LAW § 125.25(a)(ii) (McKinney 2019) (The "discovery, knowledge or disclosure of the victim's sexual orientation, sex, gender, gender identity, gender expression or sex assigned at birth" is not a "reasonable explanation or excuse" for the defendant having acted "under the influence of extreme emotional disturbance").

55. *See* Michal Buchhandler-Raphael, *Loss of Self-Control, Dual-Process Theories, and Provocation*, 88 FORDHAM L. REV. 1815, 1865 (2020). It is not clear if, according to this proposal, the decedent's "legally wrongful act" is a sufficient condition for "adequacy," or only "one piece" of provocation's "normative dimension." *Id.*

56. *See* Nourse, *supra* note 1, at 1396 ("To 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. A strong measure of that equality can be found by asking whether the emotion reflects a wrong that the law would independently punish.").

work in tandem. A defendant who forms an intent to kill is excused pursuant to element (1) for having formed that intent if and because he lacked the capacity not to form that intent, provided he's also excused pursuant to element (2) for having formed the judgment that initiated the psychological process producing that incapacity in the first place.

3. Element (3) – Partial Defense

That leaves element (3), pursuant to which a successful plea of provocation mitigates what would otherwise be the greater offense of murder to the lesser offense of manslaughter. The provocation defense is in this sense a partial defense only, and not a complete or full defense. Any theory of provocation must therefore provide some explanation for this feature of the doctrine. Element (3) thus imposes a constraint on how elements (1) and (2) are specified such that their specification yields only a partial defense.

Start with the element (1), the excusing condition. According to the partial excuse theory, element (1) – the loss of self-control element – references or describes an incapacity: a passion-induced limit on reason's power to attend to or receive (and thus to respond to) some range of reasons not to kill. With reason incapacitated in this way, the defendant would likewise lack to a corresponding degree the capacity to form an intent to do something, in response to the wrong he judged was done to him, other than an intent to kill.

Now, the extent to which passion in fact incapacitates a particular defendant's reason is a psychological fact about the defendant. That psychological fact, moreover, is a non-categorical fact: capacities and incapacities are commonly supposed to come in degrees. A person's capacity to do this or that can be more or less, and that capacity can likewise be more or less incapacitated. Nonetheless, theories of the provocation doctrine standardly recognize only two categories of incapacity. Whatever the actual extent to which a defendant's capacity not to form an intent to kill was incapacitated, that incapacity is sorted into either the "partial-incapacity" category or the "total-incapacity" category. Theories of provocation are not alone in dichotomizing psychological realities in this manner. Other criminal law doctrines likewise carve continuous reality up into some number of artificial categories as well.[57] Such theoretical and doctrinal categories necessarily fail, more or less, to map reality perfectly.

The partial excuse theory interprets element (1) such that it sets a ceiling on the extent to which the state will recognize or acknowledge, by way of the provocation doctrine, the incapacity a defendant in fact experienced. If the incapacity a defendant experienced was partial, then the state will, setting aside the possibility of forfeiture under element (2),

57. *See, e.g.*, Moore, *supra* note 26, at 356 ("[T]he law . . . in many places . . . attach[es] a bivalent remedy on what we all know is in nature a matter of continuous variation.").

recognize the defendant's incapacity to its full extent. If the incapacity a defendant experienced was in fact complete or total, however, the partial excuse theory gives that "excess" incapacity no legal effect. The interpretation the partial excuse theory assigns to element (1) simply has no way to recognize, or afford any legal effect to, any incapacity the defendant experienced beyond a partial incapacity. As the partial excuse theory interprets its elements, the provocation doctrine is willfully blind to any such excess incapacity. A defendant who was completely incapacitated is treated as if he were only partially incapacitated.

That leaves element (2), the forfeiture condition, according to which inadequate provocation triggers a forfeiture of element (1), thereby denying the defendant access to the partial excuse instantiated in that element, whereas adequate provocation triggers no such forfeiture. Inasmuch as element (3) stipulates that the defense is only partial, and inasmuch as element (1) limits the extent to which the doctrine recognizes a defendant's incapacity (such that any incapacity beyond a partial incapacity goes without legal effect), it follows that element (2) must operate, when it does, as a full or complete forfeiture. Either the partial excuse in element (1) has full effect (when the alleged provocation is adequate), or it has no effect (when the alleged provocation is inadequate). The partial excuse theory thus secures the provocation doctrine's status as a partial defense because it combines a partial excusing condition with an "all-or-nothing" forfeiture condition.

IV. The Partial Forfeiture Theory: Reason Totally Incapacitated

The psychological process on which the partial forfeiture theory rests differs in one key respect from that on which the partial excuse theory rests. Specifically, the particular cognition that produces a reason-incapacitating passion is, according to the partial forfeiture theory, something other than a judgment of reason. Specifically, that which produces reason-incapacitating passion is not the judgment that θ, but something else. This psychological difference, as we'll see, entails different interpretations for each of the provocation doctrine's three elements, thus yielding a different theory of provocation.

A. *Psychology*

According to the partial excuse theory, that which causes the defendant to experience a reason-incapacitating passion is the defendant's *judgment* that θ: that "So-and-so seriously wronged me and it would be good to achieve vindication." This judgment resulted from the operation of the defendant's reason without any interference from passion, since passion had not yet been aroused at the time the judgment was formed. As a result of this judgment, the defendant's will then forms the intent to seek vindication. Again, no passion has yet been aroused when the will does this work.

The intent to seek vindication, resulting from a movement of the will, then produces passion, and the force of this passion then partially incapacitates the defendant's reason.[58] Thanks to passion, the defendant's reason is rendered partially incapable of forming any judgment other than the judgment that killing should be chosen as the means to achieve vindication. This judgment then elicits from the will an intent to kill. The partial excuse theory, in short, starts in reason, with the formation of the judgment that θ, and ends in the will's formation of an intent to kill, with a reason-incapacitating passion arising between the two.

The partial forfeiture theory starts elsewhere. The partial excuse theory presupposes a psychology in which reason-incapacitating passion is the eventual result of reason having formed the judgment that θ. In contrast, the partial forfeiture theory presupposes a psychology in which reason-incapacitating passion results, not from the exercise or movement of reason, but from the exercise or movement of some other human power. This other power, like reason, forms cognitions, but the cognitions it forms are cognitions different and distinct in kind from the cognitions reason forms.

Call this other power the "sensory apprehension."[59] Again, this power, as I suppose it to be, does not produce judgments. It produces cognitions of a different kind or order: images, imaginings, memories, associations, and so forth, all of which I lump together under the generic heading "perceptions."[60] For present analytical purposes, a sharp contrast

58. *See supra* note 30 (discussing "overflow").

59. The "sensory apprehension" is itself, according to Thomistic psychology, a collection of other "internal" sense powers, together with the "external" senses of sight, hearing and so on. *See, e.g.*, Jensen, *supra* note 27, at 54–63.

60. Inasmuch as it rests on a psychology according to which passions result from perceptions, the partial forfeiture theory bears some relationship to so-called "perceptual" theories of emotion. *See, e.g.*, Christine Tappolet, Emotions, Values, and Agency 15 (2016) ("According to the Perceptual Theory, emotions are, in essence, perceptual experiences of evaluative properties."); Brandon Yip, *Emotion as High-level Perception*, 199 Synthese 199, 199 (2021) (defending "an account of emotion as high-level perception'). These theories take pains to distinguish themselves from so-called "cognitive" (or "judgmental") theories of emotion, although it should be noted that which theory of emotion – perceptual v. judgmental – is the "better" theory continues to be debated in the philosophical literature. *See, e.g.*, Giulio Sacco, *Recalcitrant Emotions: The Problems of Perceptual Theories*, ___ Ratio (Forthcoming 2024) (manuscript at 1) (on file with author) (concluding that the "analogy between emotions and perception is flawed and the that the appeal of these theories lies on a surreptitious shift in the meaning of the term 'perceive.'"). Perceptual theories conceptualize "emotions" either as partly constituted by "perceptions of values" or as something caused by such perceptions, whereas cognitive (or judgmental) theories conceive of emotions as partly constituted by "judgments of values" or as something caused by such judgments. *See generally* Julien A. Deonna & Fabrice Teroni, The Emotions: A Philosophical Introduction 52–75 (2012) (discussing "perceptual theories of emotion").

On my reading, the distinction between the "judgmental" theory of emotion, and the "perceptual" theory, which seems well-established among philosophers

is drawn between judgments, on the one hand, and perceptions, on the other. Reality is doubtless more nuanced and complicated. Indeed, the relationship between the power of reason (the "rational apprehension") and the power of "sensory apprehension" is complicated, confusing, and controversial.[61] Still, the distinction between judgment and perception is central to the difference between the partial excuse theory and the partial forfeiture theory, and some such distinction presumably exists.

What, then, is the difference between a "judgment" and a "perception"? More specifically, because a provoked defendant has either judged

---

who write about emotion, has gone largely, if not entirely, unrecognized in the criminal law literature on provocation. Indeed, sometimes an author will vacillate between language associated with a judgmental theory and language associated with a perceptual theory, without noticing that something important might turn on the difference. *Compare, e.g.*, Kahan & Nussbaum, *supra* note 27, at 286 (speaking about the "value *perceived*" in an "object") (emphasis added), *with id.*, at 293 (stating that the "cognitive attitudes" that are "*constituent parts* of the emotion" are "usually, *beliefs or judgments*") (emphasis added). It seems to me that Thomistic psychology has the advantage of being able to accommodate both perceptual and judgmental theories, inasmuch as it recognizes that passion sometimes results, albeit indirectly, from the operation of reason (resulting in "consequent" passions, which would be associated with judgmental theories), and sometimes directly from the operation of the sensory apprehension (resulting in "antecedent" passions, which would be associated with perceptual theories).

61. The basic problem can perhaps be stated most abstractly as the problem of the relationship between reason and sense, or between the rational part of the soul and the sensory part of the soul. On the one hand, the thesis that perception (or the operations of the sensory apprehension more generally) can causally influence the judgments a person forms (or the operations of reason more generally) would seem not to be very controversial. What one "sees" causally influences, at least *prima facie*, what one believes, as in the phrase "seeing is believing."

On the other hand, the thesis that reason can influence perception, such that what one believes can causally influence what one "sees" is probably more controversial. In the non-Thomist philosophical literature, this question appears to be framed in terms of whether or not reason can "cognitively penetrate" the senses, or what the nature is of the "border" between "cognition" and "perception." *See, e.g.*, Samuel Clarke & Jacob Beck, *Border Disputes: Recent Debates along the Perception-Cognition Border*, Phil Compass 1 (2022); Robert Cowan, *Cognitive Penetrability and Ethical Perception*, 6 Rev. Phil. Psych. 665 (2015); Dustin Stokes, *Cognitive Penetrability of Perception*, 8 Phil. Compass 646 (2013). According to some commentators, Aquinas maintained that judgment could, at least sometimes, influence perception. *See, e.g.*, Dominik Perler, *Rational Seeing: Thomas Aquinas on Human Perception*, *in* Medieval Perceptual Puzzles: Theories of Sense Perception in the 13th and 14th Centuries 213, 233–34 (Elena Băltuţă ed. 2019) (stating that Aquinas "subscribes" to the "moderate version" of the "transformative theory" articulated by Boyle); Candice Vogler, *The Intellectual Animal*, 100 New Blackfriars 663 (2019) (same); *see also* Matthew Boyle, *Additive Theories of Rationality: A Critique*, 24 Eur. J. Phil. 527, 531–532 (2016) (citing Aquinas as a proponent of the "transformative theory" according to which the sense power of human beings *qua* rational animal are "realized" in ways "distinct" from the ways in which they are "realized" in non-rational animals).

or perceived that he's been in some sense morally wronged, what's the difference between a "moral judgment" and a "moral perception"? That, it turns out, is not an easy question. Judgments of any sort are presumably propositional and conceptual, whereas perceptions may be neither. Or, perhaps perceptions are, at least some of the time, conceptual in some sense, but not propositional.[62] Or, whereas judgments are formed through reason in response to reasons for belief, epistemic reasons, practical reasons, evidence, and so on, perceptions are formed through the sensory apprehension in response to properties, including, perhaps, moral properties.[63] Clearly and sharply distinguishing judgments from perceptions is difficult, in no small part because the way in which reason and sensory apprehension interact with one another is complex.

Still, whatever the distinction between judgment and perception amounts to, the partial forfeiture theory assumes some such distinction exists. Perceptions, whatever they are, are not judgments. However, and more importantly for present purposes, perceptions, like judgments, can and do arouse passions. Indeed, perceptions cause passions *directly*, whereas judgments cause passions only *indirectly*, with movements of the will mediating between reason and passion. A powerful perception-induced passion, just like a powerful judgment-induced passion, can incapacitate reason, causing reason not to receive or attend to a range of reasons to which it would otherwise have been receptive or to which it would otherwise have attended. In other words, passion can cause a loss of self-control qua incapacitation whether passion is aroused through reason and judgment, or through sensory apprehension and perception. For convenience's sake, I'll call whatever perceptual experience produces a reason-incapacitating passion the perception that θ.

According to the partial forfeiture theory, a perception-caused passion is that which incapacitates (to some as-yet-unspecified degree) a

62. *See* Murphy, *supra* note 31, at 169 ("[S]ensory appetitive movements or passions will have to be inclinations to *non-conceptualized* primitive sets of sensory properties . . . while inclinations to objects that involve any conceptualization will have to be inclinations of the intellective appetite—the will.") (emphasis added); *cf.* Tappolet, *supra* note 60, at 18 ("If emotions are *non-conceptual* representations of evaluative properties, then it should be expected that emotions are like sensory experiences in that they allow us to be aware of certain features of the world.") (emphasis added).

63. According to one commentator, the sensory apprehension (more specifically, the cogitative power) is a power of "moral perception," but only when in some complicated way its operations are "integrated" into the judgments of practical reason. *See* De Haan, *supra* note 30, at 317. Contemporary debates and controversies surrounding the idea of "moral perception," not to mention the relationship between "moral perception" and "moral judgment," are beyond my present competence. For brief introductions to the non-Thomistic philosophical literature on "moral perception," see, for example, Anna Bergqvist & Robert Cowan, *Introduction*, *in* Evaluative Perception (Anna Bergqvist & Robert Cowan eds., 2018); James Hutton, *Moral Experience: Perception or Emotion?*, 132 Ethics 570 (2022) (arguing that "moral experience" consists of "emotions"); *see also* Preston J. Werner, *Moral Perception*, 15 Phil. Compass 1 (2020).

defendant's reason. Reason's incapacity does not mean reason lacks the power to form judgments.[64] Nor does it mean the will lacks the power to form intentions in response to those judgments.[65] Indeed, although incapacitated, in the sense that some range of reason is excluded from reason's purview, reason nonetheless forms the same judgments as does the defendant depicted in the partial excuse theory: vindication would be a good end to achieve because the decedent seriously wronged the defendant, and death would be a good means to achieve that end. This last judgment then moves the defendant's will to form an intent to kill. The key difference between the psychology of the partial excuse defendant and that of the partial forfeiture defendant is that the partial forfeiture defendant's reason is incapacitated before it forms *any* of the judgments eventually producing an intent to kill.

The partial forfeiture theory rests on a particular description of the psychological process by which a person forms a cognition that sets in motion a series of psychological events terminating in the formation of an intent to kill. One might ask what sorts of other facts are likely to be true of such a defendant, or from what sorts of other facts might one fairly infer that a particular defendant formed an intent to kill in the way the partial forfeiture theory supposes. Keeping in mind that what's important to the theory is the psychology on which it rests and not facts from which the existence of that psychology might be inferred, I am inclined to think that

64. Likewise, the partial excuse theory assumes reason's (partial) incapacity does not prevent reason from forming the judgment that killing would be a good means to achieve the end of vindication.

65. The two preceding sentences assume that the movements of reason and will, which produce judgments and intentions, are necessary to produce any human action, even when passion has "incapacitated" reason in the manner described in the text. Having said that, King cites a passage from the *Summa* suggesting that passion can indeed move the body without intervening movements of reason and will, such that the bodily movement is an "action of a human" but not a "human action." Peter King, *Aquinas on the Passions*, in AQUINAS'S MORAL THEORY, *supra* note 27, at 123 & n.43 (citing SUMMA THEOLOGICA I-II.10.3). A criminal lawyer might try to capture the difference between the "action of a human" and a "human act" by describing the former as a "reflex" and the latter as a "voluntary" act, understood as a bodily movement resulting from the formation of a "volition."

If the movement of a defendant's body in any particular case was in fact merely the action of a human (a reflex) and not a human action (voluntary act), then the defendant would not have formed an intent to kill (let alone a "volition," which Aquinas calls "use," a distinctive movement of the will), and without an intent to kill no *prima facie* case of murder would be established. The partial excuse and partial forfeiture theories thus assume that the movement of the defendant's body, despite his reason having been to some degree incapacitated by passion, was a human act, and not merely the act of a human. Horder seems to suppose that, if a defendant's reason has been "incapacitated" by passion, then reason necessarily lacks the power to move the will to form intentions. *Cf.* HORDER, *supra* note 11, at 119. I do not see why that should be so. For purposes of the partial forfeiture theory, passion incapacitates reason without reducing the resulting bodily movement to a mere "act of a human."

a defendant who forms an intent to in the way the partial forfeiture theory describes is likely to have experienced a "sudden" heat of passion, i.e., one in whom passion arises immediately and vehemently upon perceiving the allegedly provoking event; or is likely to report after the event not remembering what happened or not knowing why he killed.

B. *Doctrine*

Turning to doctrine, what difference does the partial forfeiture theory's psychology make to the provocation doctrine? If a defendant in fact lost self-control in the way the partial forfeiture theory describes, what would that psychology entail for the way in which the provocation doctrine's elements are to be understood or interpreted?

1. Element (1) – Loss of Self-Control

The partial forfeiture theory interprets the loss of self-control element in much the same way as does the partial excuse theory. A person "loses self-control" when passion limits to some degree the range of reasons available to the defendant's reason. Passion prevents reason from doing one of the things reason is supposed to do; namely, to respond to the full range of available reasons for and against the pursuit of a proposed end. The loss of self-control element in the partial forfeiture theory nonetheless differs in two ways from the loss of self-control element in the partial excuse theory.

First, it differs in relation to the point in time at which passion incapacitates reason and thus it differs in the range of judgments that result from passion-incapacitated reason. For a defendant who loses self-control in the way the partial excuse theory supposes, passion excludes practical reasons not to judge that killing should be chosen as the means to vindication, but it does not exclude any reasons available to the defendant to judge that he had been seriously wronged, or that he should choose to vindicate that wrong. For the partial excuse theory, passion does not exclude reasons related to the formation of these latter two judgments—the seriousness of the wrong, and what should be done in response to that wrong—because passion arises only after these judgments have been made; indeed, passion arises in response to these judgments.

In contrast, for a defendant who loses self-control in the way the partial forfeiture theory supposes, passion's exclusionary effect sweeps more broadly. It excludes reasons bearing on all the judgments necessary to form the practical syllogism ending with the judgment that killing is to be chosen as the means to achieve vindication. That's because passion arises antecedent to any operation of reason, not consequent to it. For that reason, the scope of passion's disabling effect on reason is more extensive than it is in the partial excuse theory.

Second, for reasons explained more fully below in Part IV.B.3, the psychology associated with the partial excuse theory differs from that of the partial forfeiture theory in relation to the degree to which passion

disables or incapacitates reason. The partial excuse theory assumes the defendant's reason was only partially incapacitated, even if it was in fact incapacitated to a greater extent. If passion in fact totally or completely incapacitated the defendant's reason, the structure of the provocation defense nonetheless forces the partial excuse theory to ignore the full extent of the defendant's incapacity. Because the partial excuse theory assumes that the provocation doctrine's excusing condition is limited to a partial incapacity of reason, it lacks the resources to recognize any incapacity greater than that.

In contrast, the psychology on which the partial forfeiture theory rests requires that passion totally or completely disable the defendant's reason, rendering reason totally or completely powerless to access any of the reasons to which it would otherwise have had access, and to which it needed access in order not to form an intent to kill. Thus, according to the partial forfeiture theory, element (1) requires a total or complete loss of self-control. How this requirement of total or complete loss of self-control is reconciled with the fact that the provocation defense is, pursuant to element (3), a partial defense, and not a complete defense, will come from the way in which the partial forfeiture theory interprets the forfeiture condition reflected in element (2).

In sum, the psychology on which the partial forfeiture theory rests is such that passion disables reason from the start, before reason forms any of the judgments eventually eliciting from the will an intent to kill; moreover, the disabling effect of passion in the psychology on which the partial forfeiture theory rests is greater in degree compared to its disabling effect in the psychology on which the partial excuse theory rests.

2. Element (2) – Adequate/Inadequate Provocation

According to the partial excuse theory, element (2) is a doctrinal category by and through which the state judges the exercise of the defendant's reason, asking if the defendant is to be excused or not for having formed the judgment that θ, a judgment he was obligated not to have formed inasmuch as it produced a reason-incapacitating passion. When the state recognizes an alleged provocation to have been adequate, it thereby excuses the defendant for having formed that judgment, despite its having produced so powerful a movement of the will as to produce in turn a reason-incapacitating passion. Conversely, when the state refuses to recognize an alleged provocation as adequate, it thereby refuses to recognize any such excuse.

The partial forfeiture theory gives element (2) a different interpretation. According to the partial forfeiture theory, the reason-incapacitating passion the defendant experienced arose from the perception that θ, not from the judgment that θ. By the time a partial forfeiture defendant forms the judgment that θ, passion has already incapacitated his reason. As such, it would be odd for the partial forfeiture theory to interpret element (2) as a doctrinal category for excusing or not a defendant for

forming the judgment that θ, since that judgment was the result of an already-incapacitated reason. Instead, element (2), according to the partial forfeiture theory, is a doctrinal category by and through which the state passes judgment, not on the defendant's exercise of reason, but on the exercise of his sensory apprehension: the capacity by which a person perceives or sees the world, including what might be called the "moral world" in which a person perceives, among other things, the actions of another as having wronged him

The way in which a person sees or perceives the moral world, and in particular the ways in which he perceives things that arouse anger, would seem to depend on how his sensory apprehension has been *habituated* to see or perceive the moral world. Judgments are reason's responses to reasons for believing or acting one way rather than another. Perceptions, in contrast, are the sensory apprehension's habitual responses to particular things in the world, including "things" in the moral world. Some people have been habituated to see some things in the moral world as wrongs; others have been habituated to see the same things as not-wrongs. So too, different people have been habituated to see the same thing as wrongs of different orders or magnitudes. What does the claim that a person's perceptions of the moral world are the result of perceptual habits entail for the way in which the partial forfeiture theory interprets element (2)?

First, element (2), according to the partial forfeiture theory, is a doctrinal category by and through which the state passes judgment on the habitual way in which a defendant perceives the moral world. When the state finds or declares an alleged provocation to have been adequate or inadequate, that finding constitutes a judgment on the defendant's habitual way of perceiving the moral world. The defendant's habitual way of perceiving the moral world is that which the state subjects to judgment by and through element (2) because the way in which the defendant perceived the moral world is, according to the psychology on which the partial forfeiture theory rests, that which produced a reason-incapacitating passion in the first place.

Second, the partial forfeiture theory's rendering of element (2) presupposes that anyone subject to the state's jurisdiction is obligated or bound to habitually perceive the moral world such that what he perceives on any particular occasion does not cause him to experience a reason-incapacitating passion. Thus, in all cases in which a defendant in fact experiences a reason-incapacitating passion he has on that occasion perceived the moral world, according to the partial forfeiture theory, in a way the state obligated him not to have seen or perceived it. In other words, any defendant pleading provocation has on the present occasion perceived the moral world in a way the state obligates him not to have perceived it.

Third, and here's the main point, an alleged provocation is adequate, says the partial forfeiture theory, if and because the state takes itself, in light of the nature of the alleged provocation, to have sufficient

reason to believe that the defendant is sufficiently habituated to perceive the moral world so as not to experience a reason-incapacitating passion. In other words, an alleged provocation is adequate if and because the state believes that the provocation was such that even someone who was sufficiently habituated to perceive the moral world so as not to experience such a passion might nonetheless, on the present occasion and in the face of the alleged provocation, have perceived it in such a way as to suffer or experience a reason-incapacitating passion as a result.[66]

66. The use of the word "habit" in the text, and not the word "virtue," is intentional. The problem, as I currently see it, amounts to this. All virtues are habits, but not all habits are virtues. Thomistic psychology posits that the cardinal virtues inhere in, or are located in, reason (where the intellectual virtue of "prudence" is located); the will (where the moral virtue of "justice" is located); and the sensory appetite (where the moral virtues of "temperance" and "courage" are located). But what about the sensory apprehension, and in particular in the cogitative power? Is any virtue located in it?
So far as one can tell, the "cogitative power" (part of the "sensory apprehension") is sometimes "integrated into the operations of practical reason" and sometimes it is not. *See, e.g.*, De Haan, *supra* note 30, at 317 (distinguishing between "antecedent and consequent acts of cogitative perception"); Dugandzic, *supra* note 30, at 139 ("[T]he cogitative power can act independently of universal reason, but it ought not to."). When the movements of the cogitative power do not "participate" in reason, the perceptions a defendant thereby experiences can be subject only to habits. *See, e.g.*, Robert C. Miner, *Aquinas on* Habitus, *in* A History of Habit 67, 73–74 (Tom Sparrow & Adam Hutchinson eds., 2103) ("[T]he imagination, memory, and the cogitative power . . . can receive habits. That not only the sensory appetite, but also the interior powers of sensory apprehension can be habituated is crucial. Were this not the case, there would be no possibility of educating the passions."). However, when the movements of the cogitative power do "participate" in reason, the perceptions a defendant thereby experiences can, albeit indirectly, "participate" in the virtue of prudence (assuming that virtue has been acquired). However, according to Pasnau, "There is no such this as having virtuous senses." Robert Pasnau, Thomas Aquinas on Human Nature 17 (2002). If so, then the psychology associated with the partial forfeiture theory must assume that the perceptions a defendant experienced were experienced without the cogitative power having "participated" in reason. Having said that, one might nonetheless ask how the cardinal virtues would figure into the psychology associated with the partial forfeiture theory. I am not entirely sure. My provisional sense is that any defendant who experiences a passion sufficiently vehement to completely incapacitate reason is necessarily someone insufficiently habituated in at least one of the cardinal virtues; and, if the unity of the virtues thesis is true, then he would be insufficiently habituated in all of them. That's because these virtues, or at least the virtue of temperance, "impedes vehement antecedent passions." De Haan, *supra* note 30, at 295. If so, then any defendant who satisfies element (1), as the partial forfeiture theory interprets that element, is necessarily someone insufficiently habituated in one or all of the cardinal virtues: the only defendants who need to plead provocation are thus necessarily "lacking in virtue," whether the provocation to which they responded is found, as a matter of law, to be "adequate" or "inadequate." The partial forfeiture theory then sorts this collection of unvirtuous defendants into, on the one hand, those who get the mitigation because the state judges that, despite their lack of virtue, they are nonetheless sufficiently habituated to see

Conversely, an alleged provocation is inadequate, says the partial forfeiture theory, if and because the state takes itself, in light of the nature of the alleged provocation, to have insufficient reason to believe that the defendant is sufficiently habituated to perceive the moral world so as not to experience a reason-incapacitating passion. In other words, an alleged provocation is inadequate if and because the state believes that the alleged provocation was such that no one who was sufficiently habituated to perceive the world without experiencing such a passion would, even in the face of the alleged provocation, have perceived it in a such a way as to suffer or experience a reason-incapacitating passion as a result.[67]

Fourth, a defendant who the state judges to be insufficiently habituated to perceive the moral world so as not to experience reason-incapacitating passion is said to be "at fault," "culpable," "blameworthy," and so forth for experiencing the reason-incapacitating passion he in fact experienced on the present occasion. Conversely, a defendant who the state judges to be sufficiently habituated to perceive the moral world so as not to experience reason-incapacitating passion is said not to be "at fault" and so on for experiencing the reason-incapacitating passion he in fact experienced on the present occasion. As explained in more detail in Part IV.B.3, this judgment—that the defendant either was or was not in this sense "at fault" for experiencing the reason-incapacitating passion he in fact experienced—in turn governs the extent to which the state recognizes the resulting loss of self-control as an excuse.

In short, one might roughly say that element (2) is an imperfect proxy by which the state sorts defendants, all of whom experienced a reason-incapacitating passion, into those who it believes are habitually "good" perceivers of the "moral world" and those who are not; or in other words, into those the state believes tend to perceive the "moral world" in an "appropriate" way and those that do not. Element (2) is only a rough way for the state to sort defendants into those it believes are habitually good moral perceivers from those who it does not because the state's judgment one way or the other is based only on the way in which the defendant perceived the moral world on the present occasion, which may not by itself be reliable evidence of the way in which he in fact habitually perceives it.

---

or perceive the "moral" world as the state obligates them to see or perceive it; and, on the other, those who do not get the mitigation because, in addition to their lack of virtue, they are also insufficiently habituated to see or perceive the "moral" world as the state obligates them to see or perceive it.

67. According to one commentator, the habits by which a person perceives the world are "extremely difficult," though not impossible, to change:

> Aquinas explains that the estimations [associations, imaginings, etc.] stored in the memorative power are literally ingrained in a person's soul. To say this is not merely metaphorical. Because estimations of the interior sensory powers require a bodily organ, they are literally "immuted" or imprinted on the soul.

Miner, *supra* note 29, at 80.

This rendering of element (2) can perhaps make some sense of something a number of commentators have found nonsensical. One fairly common test for judging the adequacy of an alleged provocation asks if "the [alleged] provocation . . . was enough to make a reasonable man do as the defendant did?"[68] That test, say some, is nonsense: whatever else may be true of a reasonable person, a reasonable person never, no matter what the alleged provocation, gets so angry as to lose self-control such that he ends up forming an intent to kill someone.

This charge of nonsense is fair, provided the reasonable person not only habitually perceives the moral world so as not to experience any reason-incapacitating passion, but is also someone in whom that habit *never* fails. The charge is less fair, however, if the reasonable person is someone who not only habitually perceives the moral world so as not to experience any reason-incapacitating passion, but also, being human, is a being in whom the requisite habit on occasion fails, as all human habits *qua* habits sometimes do. Habits habituate. They do not necessitate. If so, then it is not nonsense to say that some provocations are such that they might make even a reasonable person – understood as a habitually "good" moral perceiver – find himself so suddenly angry as to disable reason. If so, then some provocations might indeed be enough to make "a reasonable man to do as the defendant did."

3. Element 3 – Partial Defense

As with any theory of provocation, the partial forfeiture theory needs to explain why the provocation defense is a partial defense only, and why that partial defense is available only when the alleged provocation is judged to have been adequate. The partial forfeiture theory's explanation, like the partial excuse theory's, results from a further specification of elements (1) and (2).

According to the partial excuse theory, the loss of self-control to which element (1) refers is established if the defendant's reason was, as a result of passion, at least partially incapacitated, such that he partially lacked the capacity not to form an intent to kill. Element (2) is in turn interpreted as an all-or-nothing forfeiture condition. If an alleged provocation is inadequate, the partial excuse in element (1) gets no legal recognition: it's lost completely. If an alleged provocation is adequate, the partial excuse in element (1) is given full legal recognition: it's not lost at all.

The partial forfeiture theory flips this around. According to the partial forfeiture theory, the loss of self-control in element (1) is not established unless the defendant's reason was, as a result of passion, completely or fully incapacitated, such that he lacked the capacity full-stop not to form an intent to kill. The partial forfeiture theory thus requires the defendant to establish that passion impaired his reason to a greater extent compared to the partial excuse theory. But if the partial forfeiture theory's excusing condition is a condition that would provide a complete or total excuse,

68. Homicide Act 1957, 5 & 6 Eliz. 2 c. 11, § 3 (UK).

then why is not the provocation defense a full defense, at least when the alleged provocation is adequate? What explains why the defense is a partial defense only, even when the alleged provocation is adequate?

The answer comes from the interpretation the partial forfeiture theory gives to element (2). When an alleged provocation is judged to have been inadequate, the partial forfeiture theory says that the defendant loses or forfeits the theory's associated excusing condition completely, such that the defendant is liable for murder. When the alleged provocation is inadequate, element (2) thus functions the same way in the partial forfeiture theory as it does in the partial excuse theory. It results in a complete forfeiture of the associated excusing condition.

When the alleged provocation is "adequate," however, element (2) functions one way in the partial excuse theory and in a different way in the partial forfeiture theory. For the partial excuse theory, an "adequate" provocation results in no forfeiture of the associated (partial) excusing condition. In contrast, for the partial forfeiture theory, an "adequate" provocation results in a partial forfeiture of the associated (complete) excusing condition. In other words, for the partial forfeiture theory, an excusing condition that would otherwise have afforded a full defense nonetheless results in a partial defense because when the alleged provocation is adequate the excusing condition is still partially forfeited.

This partial forfeiture, despite the alleged provocation having been judged adequate, reflects competing reasons: reasons to recognize the defendant's passion-induced, complete incapacity as a full excuse and reasons not to recognize it as a full excuse. On the one hand, it reflects the state's judgment that the defendant, having formed an intent to kill only as a result of reason-incapacitating passion, could not have intended otherwise than to kill; and that the defendant is someone who, in the state's judgment, habitually perceives the moral world so as not to experience such reason-incapacitating passion. On the other hand, it also reflects the state's judgment that the defendant's habit of perceiving the moral world as the state obligates him to perceive it did in fact fail on the present occasion. Although he habitually perceives the moral world as he's obligated to perceive it, he did in fact, on the present occasion, perceive it in such a way as to experience such a reason-incapacitating passion.

So understood, the partial forfeiture theory portrays a manslaughter conviction as the price a defendant pays for having perceived the moral world on the present occasion such that he experienced a passion blinding him completely to reason and forming an intent to kill as a result. That price is not as high as it might otherwise have been because, given the nature of the alleged provocation, the state continues, as one might say, to trust the defendant as someone who habitually perceives the moral world so as not to experience such passion. A murder conviction, in contrast, is the price a defendant pays for having lost that trust.

The provocation defense has been said to be a concession to human frailty.[69] That phrase can mean many different things. According to the partial forfeiture theory, it does not mean, or does not only mean, that human beings are frail because they're vulnerable to sometimes vehement passions, though that's true enough. Rather, it means that human beings are frail because human habits can be frail. Even someone who habitually perceives the moral world as the state demands he perceive it might sometimes, in the face of some provocation, misperceive it. Even someone lucky enough to have been habituated to perceive the moral world so as not to experience reason-incapacitating passions might nonetheless find himself on occasion vulnerable to such passion.

Conclusion

I doubt that many defendants will fit the psychological profile on which the partial forfeiture theory rests. I would guess that defendants who form an intent to kill because passion completely blinds their reason to reasons not to kill, leaving them powerless to have intended otherwise than to kill, are not many. But I also doubt that the set of such defendants is an empty set, and among defendants within the set, some will go to prison for murder; others for manslaughter.

Those sent to prison for manslaughter, but not murder, are sent there because they did, after all, form an intent to kill. At the same time, they could not have intended otherwise, given the reason-incapacitating passion they experienced. The state, moreover, continues to have faith in them as someone who habitually perceives the moral world so as not to experience such passion. Those sent to prison for murder are sent there because they too, after all, formed an intent to kill. Likewise, they too could not at the time have intended otherwise, given the reason-incapacitating passion they experienced. But, unlike those reason-incapacitated defendants who get sent to prison for manslaughter, those who get sent to prison for murder get sent there, not just because they misperceived the moral world on the present occasion, but also because the state no longer has faith in them as someone who habitually perceives the moral world so as not to experience such passion.[70]

69. Dressler, *supra* note 2, at 973 ("We must remember that the provocation defense is based to a considerable extent on the law's concession to ordinary human frailty . . . .").

70. When the state judges that a defendant was "insufficiently habituated" in the manner described in the text, is it necessarily making a judgment about the defendant's "character," and if so, should the time a person spends in prison depend in any way on the state's judgment about his "character"? I would not try to answer those questions here. Nonetheless, I say a little bit about provocation and "judging character" in Appendix B.

Appendix A

The generic statement of the doctrine of provocation given in the text will probably raise questions and eyebrows. Here are some additional comments bearing on why that statement is formulated as it is.

1) The generic statement's elements are presented as an affirmative defense, not as elements of a substantive crime labeled "manslaughter." That difference in legal form does not, for present purposes, make any difference in substance. The main difference this difference in legal form makes—between the doctrine of provocation as an element of a crime denominated manslaughter and as an affirmative defense—is procedural: to allocate the burdens of production and persuasion to one party or the other.

2) Because element (1) of the generic statement presupposes the defendant formed an intent to kill, the text does not discuss how, if at all, anything said in the text about the two theories of provocation it describes would need to be changed if a defendant were charged with a form of murder based on a kind of culpability other than intent to kill, e.g., an intent to inflict grievous bodily harm, or on a "depraved heart."

3) As a matter of positive law, the doctrine of provocation is limited to cases in which the defendant is charged with (some form of) murder, and no other crime. The generic statement reflects this limitation. Commentators have a hard time making sense of this limitation. Some believe it makes no sense; others believe some sense can be made of it.[71] Perhaps the limitation is simply an historical artifact, for which no principled justification can be found or offered. In any event, I would not try to make sense of the murder limitation. It's therefore fair to say that neither the partial excuse theory nor the partial forfeiture theory fully explains the doctrine, inasmuch as neither tries to explain why the doctrine does not extend to offenses other than murder.

4) The generic statement states the objective condition as "adequate" provocation. It does not use the phrase "reasonable" provocation. The phrase "adequate" provocation is used to avoid any implication that the test for the objective condition must or should be stated in terms of what would cause a "reasonable person" to experience a reason-incapacitating passion or loss of self-control, and so on.

My sense is that using or relying on the idea of a reasonable person to identify what is or is not adequate provocation, although common, causes more problems than it's worth. Reliance on the reasonable person tends to cause people, or at least criminal lawyers, to start asking questions about what characteristics, facts, traits, and so on, which are true of the defendant, should be imputed to the reasonable person for purposes of judging the adequacy of an alleged provocation. That way of

71. *See, e.g.*, Horder, *supra* note 11, at 135–36 (provocation limited to murder for "practical" reasons); Berman & Farrell, *supra* note 16, at 1105 (offering "two reasons" why a defense like provocation is "unnecessary" for offenses other than murder).

thinking about what kinds of provocation the state should count as adequate, and what kinds it should count as inadequate, does not strike me as particularly helpful. But, again, I would not say anything more here.

5) The generic statement uses the word "passion" rather than the word "emotion." It does so because passion is the concept used in Thomistic psychology to describe the relevant psychological phenomenon.[72] Emotion is apparently a concept developed in the nineteenth century. According to one source, "[I]t was the secularization of psychology that gave rise to the creation and adoption of the new category of 'emotions.'"[73]

6) Because the generic statement includes loss of self-control among its elements, any theory built around or on that statement will necessarily portray or present the provocation doctrine, in one way or another, as an excusing doctrine. That is, if and when a defendant receives the benefit of the doctrine, he receives that benefit because the state makes his choice to kill (i.e., his formation of an intent to kill) excusable, or at least not punishable, to some extent or another, even if some other condition must also be established before he receives that benefit. As such, the two theories discussed in the text will not apply to any statutory or common-law formulation of the doctrine that does not, in some way or another, include something amounting to loss of self-control among its elements.

7) The generic statement makes no mention of mistakes of fact. So far as I know, the law in most or all jurisdictions, whether by way of statutory language or caselaw, makes some provision for how to deal with instances in which the defendant makes a mistake about the facts alleged to constitute adequate provocation, i.e., the facts upon which the state determines if the alleged provocation was adequate. The common law, for example, is said to require the adequacy of an alleged provocation to be judged based on the facts as the defendant reasonably believed them to be, whereas the Model Penal Codes states that the "reasonableness of" the defendant's "explanation or excuse" for the "extreme emotional disturbance" under the influence of which he committed murder, "shall be determined . . . under the circumstances as he believes them to be."[74]

For present purposes, I set aside how the doctrine should be formulated to address such mistakes. It does bear mentioning, however, that such mistake provisions would seem to presuppose the defendant formed some judgment as to the existence of the provocation alleged, which in turn caused the defendant to experience passion. The partial forfeiture theory, however, supposes that the thing causing the defendant to experience a reason-incapacitating passion was not a judgment at all: it was a perception. Any judgment such a defendant makes as to the nature of the provocation would of course have been the result of an exercise of

72. *See supra* note 27.

73. *See* Thomas Dixon, From Passions to Emotions: The Creation of a Secular Psychological Category 4 (2003). *See generally* R.E. Brennan, The History of Psychology: A Thomistic Reading (2019) (orig. pub. 1945).

74. Mode Penal Code § 210.3(1)(b).

reason, but that judgment would have been made at a time when passion had already incapacitated reason.

8) The generic statement says nothing about the reasonableness or not of a defendant's loss of self-control. It's nonetheless sometimes said, in both the literature and caselaw, that the doctrine of provocation requires some judgment as to the reasonableness of a defendant's loss of self-control, or the reasonableness of the defendant's level of self-control.[75] In other words, some judgment about reasonableness is sometimes thought to be required as to both the gravity of the provocation alleged and the defendant's loss of self-control as a result of having cognized that alleged provocation in some way.

So far as I can tell, however, judging the reasonableness of a defendant's loss of self-control only makes sense when the psychological process by which a defendant forms an intent to kill is the process described in the abandoned self-control account: in other words, only when the defendant lost self-control in the extended sense that he abandoned self-control. If so, then the phrase reasonable loss of self-control, insofar as it is taken to be an element of the doctrine, could function as doctrinal language by which the states judges as excusable or not a defendant's decision or choice to stop exercising self-control, given whatever dysphoria, stress and so on the defendant experienced as a result of continuing to exercise self-control.

One might, of course, believe that if the choice to stop exercising self-control will necessarily result in the formation of an intent to kill, then choosing to stop exercising self-control should never be excused, partially or otherwise. When choosing to stop exercising self-control will result in the formation of an intent to kill, that choice should, the thought goes, be inexcusable no matter how hard or difficult a defendant experienced the choice to exercise.[76] That's not to say it would be illegitimate for a state to provide some form of mitigation to such a defendant. The state has the authority, within limits, to structure its criminal law as it sees fit. Nonetheless, insofar as the doctrine of provocation includes loss of self-control as an element, as I suppose it does, and not merely the abandonment of self-control, a plausible theory of provocation needs to account for the former, not the latter.

Appendix B

The partial excuse and partial forfeiture theories each include a forfeiture condition. The forfeiture associated with the partial excuse theory is established when, in the state's judgment, the defendant inexcusably misjudged the moral world as the state obligated him to judge it. The forfeiture associated with the partial forfeiture theory, when that forfeiture is complete or total, is established when, in the state's judgment, the defendant is insufficiently habituated to see or perceive the

75. *See, e.g.*, Dressler, *supra* note 33, § 31.07[B][2][b][ii], at 522.

76. *See, e.g.*, Stephen P. Garvey, Guilty Acts, Guilty Minds 149–153 (2020) (analyzing provocation as a species of "dysphoric duress").

moral world in the way the state obligates him to see or perceive it. These judgments are made when state actors apply the adequate provocation element to the facts of a particular case.[77]

To what extent does the way in which the two theories interpret the adequate provocation element make it necessary for the state to make judgments whose object is the defendant's "character"? And, to whatever extent that might be, is any such exercise of state authority somehow, for some reason, something about which one should be troubled? Those are worthwhile questions, but I would not try to address them here. Among other things, answering the first question would require a much more precise analysis of the concept of character, as well as an account of the way in which a person's character figures into the choices he makes on any particular occasion. Answering the second question would require an account of the moral limits of state authority in connection with defining and administering the offenses and defenses constituting the substantive criminal law.

Having said that, consider, for example, what Dan Kahan and Martha Nussbaum say, in their influential article, about the provocation doctrine and judging character. Start with the familiar claim that, as a matter of principle, the state should not (and according to some, presently does not) put itself in the business of judging character as it goes about defining and administering the substantive criminal law's offenses and defenses. Criminal liability, so the claim goes, should depend on the choices a person makes, not on any judgment about the content of his character.

In response to the claim that a state, or at least a "liberal" state, should not be in the business of judging character when it formulates rules for criminal liability and goes about applying those rules to particular cases, Kahan and Nussbaum allege that "it simply could not be otherwise."[78] In other words, they say, the existence of something called the criminal law somehow necessarily requires the state to make judgments about a defendant's character when it assigns criminal liability. A criminal law not requiring any such judgment is a conceptual impossibility: the criminal law necessarily entails judgments of character in some way or another. Of course, some criminal law writers would probably beg to differ, but suppose we take the allegation as true.

With respect to the provocation doctrine in particular, Kahan and Nussbaum believe that whatever mitigation the provocation doctrine affords should not be extended to defendants who have failed to "shape their characters . . . in accordance with prevailing norms of reasonableness."[79] Defendants with characters not shaped in accordance with "prevailing norms of reasonableness" are, presumably, defendants

77. Of course, state actors in the business of applying the "adequate" provocation element to the facts of a particular case may not realize they're making judgments of the kind the two theories say they're making. But that, according to the two theories, is what they're doing, whether they're aware of it or not.

78. Kahan & Nussbaum, *supra* note 27, at 360.

79. *Id.* at 366.

who, in that sense at least, have "bad" characters. If so, then according to Kahan and Nussbaum, defendants with "bad" characters who form an intent to kill in the heat of passion, should be convicted of murder, not manslaughter. But what, they ask, if a defendant with a "bad" character, who's thus convicted of murder, "did not have the degree of control over his character development that we usually do,"[80] whatever having the "usual[]" degree of control amounts to?

Bracketing the many questions one could ask about the relationship between character, responsibility, control, and criminal law, Kahan and Nussbaum say in response that, all else being equal, the state should consider, albeit only at the sentencing stage, being merciful toward such a defendant, i.e., a provoked defendant who's been convicted of murder, who has a bad character, but who lacked the usual degree of control over the formation of his character.

However, all else is not equal, because, they say, the state *should not* be merciful to all defendants who fall into this group. Specifically, the state should not be merciful toward a defendant with a bad character, despite his not having had the usual degree of control over its formation, if the nature or content of his bad character was "illiberal," in the sense in which Kahan and Nussbaum use the word "illiberal." Kahan and Nussbaum give two examples of such illiberal characters: those whose characters are described as "racist," and those whose characters are described as "homophobic." In the case of a defendant whose bad character fits a description such as those, the state should not be merciful, they say, even if the defendant lacked the usual degree of control over the formation of his character. Why is that?

Because, say Kahan and Nussbaum, mercy should be shown to defendants with bad characters, but only if: 1) they lacked the usual degree of control over the formation of their characters, *and*; 2) being merciful would "supplement[] and enrich[] the disposition of [the defendant's] particular case."[81] Showing mercy to a defendant whose character was not only bad, but bad-because-illiberal, could, depending on the facts, satisfy condition (1), but, according to Kahan and Nussbaum, it might not satisfy condition (2): it might not "supplement[] and enrich[] the disposition in the particular case." On the contrary, showing mercy to bad-because-illiberal characters "might [actually] . . . impoverish the statement made by conviction, *even assuming that the offender's unfortunate upbringing made an essential contribution to his crime*."[82]

The reasons why showing mercy to bad-because-illiberal characters might "impoverish the statement made by conviction" are, so far as one can tell, consequentialist in nature: showing mercy to "bad-because-illiberal" characters might, as Kahan and Nussbaum say, send the wrong

80. *Id.*
81. *Id.* at 370.
82. *Id.* (emphasis added).

"message of deterrence,"[83] *i.e.*, a "message that fosters and gives comfort to racism and homophobia and other reprehensible feelings," even when the defendant lacked the usual degree of control over the formation of his character.[84] Isupran short, it might not be appropriate for the state to show mercy to someone with a bad-because-illiberal character, no matter how unfortunate the upbringing that made an "essential contribution to his crime," because doing so would send the wrong message of deterrence.

83. *Id.* at 360.
84. *Id.*