A Critical Introduction to the Symposium

Kyron Huigens

Benjamin N. Cardozo School of Law

Follow this and additional works at: https://repository.law.umich.edu/mjlr

Part of the Criminal Law Commons

Recommended Citation

Available at: https://repository.law.umich.edu/mjlr/vol43/iss1/2

https://doi.org/10.36646/mjlr.43.1.critical

This Introduction and Commentary is brought to you for free and open access by the University of Michigan Journal of Law Reform at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in University of Michigan Journal of Law Reform by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
A CRITICAL INTRODUCTION TO THE SYMPOSIUM

Kyron Huigens*

I. APPROACHES TO THE PROBLEM

Ambitious. No other word seems to capture Reid Fontaine’s *Adequate (Non)Provocation and Heat of Passion as Excuse Not Justification*. This might explain the quality of the commentary in this Symposium and the prestige of the commentators. There is no doubt about either one. The impression the commentary gives, from beginning to end, in essay after essay, is that the authors found Fontaine’s article to be engaging, provocative, inspiring where it is right, and interesting where it is wrong. All of the participants express their thanks for the opportunity to participate. I will express my special gratitude for the opportunity to read all the essays together in advance of their publication. Some of the participants mention their initial reactions to being asked to participate in this Symposium. I must admit that I declined the offer initially because I was at work on my own article on provocation. However, when Fontaine asked me to write an Introduction and told me who had submitted essays, I leapt at the chance. I expected that this group would be helpful to my own thinking about provocation, and it is an understatement to say that things have turned out this way. So much so, in fact, that my simple Introduction sprawled into the contribution to the Symposium that I originally declined to produce. What made the difference, of course, was the opportunity to consider and in some cases argue against the other contributions. If my critical comments miss the mark, I hope I will at least have drawn attention to some of the best parts of these essays.

The debate over how to characterize the partial defense of heat of passion or provocation is both settled and unsettled. Those who characterize the mitigation as an excuse far outnumber those who characterize it as a justification, but there is little if any agreement among those in the “excuse” camp on why this is the right description. Fontaine addresses both the distinction and the description, and claims to show definitively that the partial defense of heat of passion is an excuse, not a justification.

*Professor of Law, Benjamin N. Cardozo School of Law.

Fontaine argues that the law recognizes four kinds of cases as provocation, in each of which no actual provocation appears as an objective matter. These cases of "adequate non-provocation" include: (1) mistakes about the circumstances constituting the supposed provocation that are (a) reasonable or (b) unreasonable; and (2) otherwise good cases of provocation that involve the killing of a third party (a) accidentally or (b) intentionally. Fontaine's argument is that these kinds of cases have been recognized as good cases of provocation, but that they cannot be recognized as good cases of justification because justification requires the vindication of true rights or the prevention of real wrongs. He concludes that provocation therefore cannot be a justification.

Fontaine reinforces his argument with some recent work in the field of psychology on the relationship between emotions and violence. Some violence is instrumental, but cases of provocation are instances of reactive violence. This description denotes an independence from the truth of the matter in cases of provocation that makes them fit well with Fontaine's description of provocation as a subjective defense. He appears to take this as further confirmation of his thesis that provocation cannot be a justification because justification supposes real rights, wrongs, and harms, instead of mere perceptions of these things.

Before previewing the essays that respond to these arguments by Professor Fontaine, I want to note one way of reading Fontaine's argument that would make it clearly wrong, because, unfortunately, his framing of the issue suggests that he is advancing just such an argument. The question of whether provocation can or cannot be a justification might be framed as asking whether there are cases in which an act that qualifies for the provocation mitigation is also an act that qualifies for a justification defense. If so, then provocation can be a justification.

The problem with this framing is that it commits what I call the extensional fallacy—the primary symptom of which is to talk in terms of "provocation cases" and "justification cases" instead of in terms of provocation as an argument and justification as an argument. On a given set of facts a provocation argument might be made, a justification argument might be made, or both arguments might be made. What, then, does it mean to ask whether provocation can be a justification? Is the answer to this question "no" if and only if provocation and justification never apply to the same case—that is, if their extensions do not overlap? If this is the right framing, then the answer must be "yes, of course provocation can be a justification." An evil person is about to throw my child off a
cliff, so I kill him in a fit of rage, thus saving my child’s life. Both arguments can be made in such a case. This implies that provocation can be a justification. Problem solved. But that cannot be right because by that logic the fact that there are cases in which one argument can be made but the other one cannot implies the contrary—that justification and provocation are mutually exclusive.

To show that provocation cannot be a justification will require more than to show that some cases of provocation are not cases of justification. The issue is the logic of the respective arguments, not the extension of each argument. Justification acquits whereas provocation mitigates, but the significance of this difference goes beyond these outcomes. There is something behind the difference, and the question is whether the difference and the thing behind it are integral or epiphenomenal to the relationship between the defenses.

I bring up the extensional fallacy because it appears that Fontaine commits it right at the outset. His argument seems to be that provocation cannot be a justification because four classes of cases support a provocation argument but not a justification argument. He goes further, however, to explain why this is so—that a good argument of provocation in the absence of real provocation makes that defense different from a justification argument, because justification concerns real rights and genuine wrongs—and this is where his critics engage him.

With that orientation, let us now take a look at the critics’ arguments.

II. Is This Worth Doing?

Gabriel Chin asks a pertinent question. Is there any practical value to the inquiry being conducted in this Symposium? His answer is no, and he makes a disturbingly strong case for it.

The imposition of punishment requires moral justification but, as Chin points out, not to impose punishment requires no such justification. Neither the distinction between the defenses nor a theoretical delineation of this distinction contributes anything to the moral justification of punishment. If the theory of punishment is normative legal theory—and this seems to be the general consensus—then to delineate the differences between excuse and justification serves no purpose.

Chin deals handily with the standard arguments for the importance of the excuse/justification distinction. For example, the distinction is commonly said to clarify the liability of accomplices, which otherwise turns on vague notions such as "community of purpose." The distinction tells us that an accomplice to a justified principal can assert the justification defense because there is no crime to be an accomplice to, whereas an accomplice to an excused actor cannot assert the excuse defense because there is a crime to be an accomplice to. Chin points out that the former accomplice does not consistently escape criminal liability. The accomplice might be guilty of felony murder regardless of the justified status of his co-felon. Or the accomplice might be guilty of the separate crime of attempting to be an accomplice to a justified actor. The excuse/justification distinction does little to clarify this more complex picture.

Chin questions the common claim that the distinction between excuse and justification is needed to clarify the law's moral message. This is not plausible, he argues, because the distinction conveys no moral information at all. It is invisible in practice. Criminal verdicts are general verdicts, and a system of special verdicts that could adequately convey the moral messages implicit in the excuse/justification distinction would be too complex to be viable. Criminal codes communicate little more than general verdicts do because, like general verdicts, codes do not label defenses according to the excuse/justification distinction.

Concerning this last point, however, I think there is a stage between the drafting of a code and the issuance of verdicts at which the distinction and its moral significance might be visible. The court's instructions to the jury lay out the terms of the defenses in detail sufficient to identify the differing rationales of excuses and justifications. Appellate litigation over jury instructions can communicate their moral messages. For example, the principal litigation in the prosecution of Bernhard Goetz, the subway shooter, concerned the proper formulation of self-defense. Did New York require proof of a reasonable belief in justifying circumstances, or merely an honest belief? The public paid attention to this litigation, and vigorously lodged its objections and suggestions. Now suppose that Dick Cheney were prosecuted for authorizing torture and that his legal advisor David Addington were alleged to be his accomplice. Cheney might argue lesser evils

and, in the alternative, mistake about lesser evils—on both the lesser evil as a matter of law and lesser evil as a matter of fact prongs. Where would this leave Addington? He would certainly insist on a set of instructions taking a position to his advantage, and the prosecution would contest them. How likely is it that the issue of Addington’s being an accomplice to Dick Cheney would not play out prominently in the public square? I cannot imagine that it would not draw at least as much attention as the Goetz case. One can hope, at least, that the debate would clarify the law’s moral message for the public.

III. PROVOCATION AND THE NATURE OF CRIMINAL RESPONSIBILITY

Samuel Pillsbury’s superb contribution to this Symposium offers a series of powerful challenges to Fontaine’s thesis, offering essential points of clarification along the way.5 The first of these points is deceptive, because it appears to muddle the basic question. Is provocation a justification or excuse? Pillsbury says yes.

The defense of justification in criminal law is a claim that one has acted correctly, and not unlawfully, on the particular occasion, even if one’s act violates a criminal prohibition. One way to describe a justified act is to say that it is reasonable, and the use of the word “reasonable” in virtually all provocation formulations suggests that provocation is a justification defense. But this is clearly unsustainable. As Pillsbury points out, the circumstances of a provoked killing cannot make this wrong a right—as evidenced by the fact that it is a mitigation, not a full defense. “One does not get sent to state prison for years and suffer all of the consequences of a serious felony conviction for doing what society considers to be a morally good, right or acceptable action.”6 In short, provocation is not a member of the class or category of justification defenses.

This is not to say, however, that provocation cannot be given a justification reading. Justification and excuse have characteristics or qualities that a partial defense might share, even if it does not qualify for classification as a justification, on an equal footing with self defense or defense of others; or for classification as an excuse, on an equal footing with insanity or minority. It is in this sense that provocation is both a justification and an excuse: it can be read as both because it has characteristics of both.

6. Id. at 145.
If the choice between excuse and justification is a question of how to characterize provocation—of whether provocation can be read as a justification—then the question is really provocation's relationship to criminal responsibility generally. Fontaine understands criminal responsibility in a way that facilitates characterizing provocation as an excuse. Pillsbury understands criminal responsibility in a way that precludes characterizing it exclusively as an excuse or as a justification. In fact, in Pillsbury's view, provocation's relationship to criminal responsibility is so uncertain that provocation itself might be considered untenable.

Pillsbury argues that not even the best interpretations of provocation add up to the negation of criminal responsibility, even in part. Merely to recognize some incapacity of the defendant is not to say the person is justified or excused. No description of a defendant, standing alone, logically prescribes or proscribes criminal liability. An is not an ought. This much is obvious, but Pillsbury adds the more subtle point that not every psychological feature is even relevant to criminal responsibility—indeed, most are not. The psychopath clearly lacks some capacities and their absence could logically support a defense to crime. Psychopathy, however, does not support a defense to crime, because these missing capacities are simply irrelevant to the normative features of criminal responsibility. Criminal responsibility depends on "the social-moral meaning of action, on the defendant's reasons for action. Although this involves an assessment of the defendant's choice, including reasons for emotion, judgment does not rely on reconstructing the internal processes leading to the defendant's decision." We will not permit the psychopath to raise a defense of insanity despite the fact that his condition otherwise amounts to an eligible mental disease or defect. We might treat the provoked actor the same way, taking his being upset to be irrelevant to his responsibility for the grave crime of murder. What matters is the bearing his reasons have on our best conception of criminal responsibility.

Pillsbury is surely right that inculpation and exculpation in provocation involve evaluating the defendant's reasons for action. Taking the analysis in this direction, however, lands it in uncertain territory. Reasons analysis has been extensively developed and refined by English legal philosophers, notably Joseph Raz\(^8\) and John

7. Id. at 143.
Gardner.\(^9\) Gardner has argued, against conventional wisdom but in line with Pillsbury's analysis, that provocation is not a claim of partial non-responsibility.\(^9\) It is not based on the impairment of reason by extreme emotion, in a way analogous to a claim of insanity. The insane actor acts against reason, but the provoked actor does not. Very much to the contrary: the provoked actor acts for reasons. The paradigmatic cuckolded husband acts as he does for the reason that he has been cuckolded. He might not act with cool deliberation, but he does act for a reason. Gardner concludes that the provoked actor is excused in the sense in which English law and theorists use that term.\(^11\) In Gardner’s analysis, he is acquitted, if at all, because of the nature of these reasons and the quality of his practical reasoning about them.

The problem is that there is no obvious analogue to English excuses in American punishment theory. The closest match to the English excuse as Gardner interprets it is the fault in wrongdoing that is constituted here as elements of an offense—the negation of which constitutes the defense of mistake of fact. (Mistake is not analyzed as a failure of proof in England because, until recently, very little of English criminal law was codified.) But if we think of provocation as involving fault in wrongdoing and the elements of offenses, then we have to take provocation at face value. That is, we will have to take it as an alternative definition of manslaughter, on a par with reckless homicide.

The overwhelming consensus is that provocation manslaughter cannot be such an offense, and it is not hard to see why. If provocation is just an alternative definition of manslaughter, then the defendant’s making a prima facie case of provocation would impose a burden of persuasion on the prosecution to prove this manslaughter beyond a reasonable doubt. But how can the prosecution, having proved a murder beyond a reasonable doubt, now be expected to prove that the defendant was provoked and that he therefore committed only manslaughter? Given that the


\(^{10}\) See John Gardner, Rationality and the Rule of Law in Offences Against the Person, 53 Cambridge L.J. 502, 509-11 (1994).

\(^{11}\) Notice that the terminology is tricky here. Gardner denies that provocation is a claim of non-responsibility akin to insanity, but says the sufficiently provoked actor is excused. \textit{See id.} This will seem inconsistent to an American reader. In England, however, a defense of non-responsibility such as insanity is termed (quite reasonably) a defense of non-responsibility. Justifications in England are the same as justifications here. The English excuses, however, include mistake of fact and duress—these being taken, like provocation, to involve acting for reasons, instead of an incapacitation of the ability to reason. If provocation is not a justification nor a claim of non-responsibility, then it lands in the category of excuses, as Gardner says, and (adjusting for terminology), as Pillsbury implies.
prosecution, murder conviction in hand, would otherwise have no incentive to prove provocation manslaughter, the murder conviction would have to be nullified. But this could only be done by stipulation, because the defendant's prima facie case of provocation would not establish a reasonable doubt about the murder. And then what if the defendant chooses to roll the dice, go for a full acquittal, and try to raise a reasonable doubt about his being provoked? If the prosecution should fail to prove that the defendant was provoked, then, having stipulated away the murder conviction, we are left with the appalling result that a defendant once fairly and truly convicted of murder would be acquitted.

This analysis is surely a worst-case scenario, and we can rest assured that Pillsbury will take the idea that provocation involves an evaluation of the provoked defendant's reasons in a different, less disastrous, direction from this. But the direction that Pillsbury points us in may have lesser hazards of this kind, and it is best to be wary of them.

IV. The Conceptual Case for Provocation as Justification

Given that Marcia Baron, alone among the participants, is a philosopher and not a law professor, it is notable that she examines the case law and treatises that Fontaine invokes more closely than anyone else, and finds his doctrinal arguments unpersuasive. In addition, she gives us what we naturally expect from a philosopher: a conceptual account of justification that easily accommodates provocation, contradicting Fontaine's claim that this cannot be done.

As mentioned at the beginning of this Introduction, Fontaine defines four classes of adequate non-provocation cases, all of which he says ought to be recognized as good claims of provocation; and all of which, he claims, cannot be framed as good cases of justification. These involve: (1) mistakes about the circumstances constituting the supposed provocation that are (a) reasonable or (b) unreasonable; and (2) mistakes consisting of killing a non-provoking third party (a) accidentally or (b) intentionally.

Baron's refutation of Fontaine on the cases falling under (1)(a) fits seamlessly with her conceptual argument. Is it impossible to treat a case of reasonable mistake about circumstances constituting provocation as a justification? Baron's answer is: "yes and no." She

makes a distinction between being materially justified and formally justified. A man who believes he is giving a sick woman deadly poison intending to speed her death, but who gives her life-saving medicine instead is materially justified. A man who believes he is giving a sick woman life-saving medicine intending to save her life, but who gives her deadly poison instead is formally justified. It should be obvious that the materially justified man is guilty of murder and that the formally justified man is not. If the best judgments of fair-minded observers differ on this question, this might be attributable to ambiguities in the idea of justification that are even more thoroughly concealed in law than they are in ordinary conversation. We can think of justification either as "\( \phi \) is justified" or as "A is justified in doing \( \phi \)." Looking at justification in the former sense will lead one to focus on material justification, while looking at justification in the latter sense will lead one to focus on formal justification. Fontaine focuses on material justification, and it is enough to refute his argument simply to point out that it is possible, and probably more common, to view the justification defenses in criminal law as a matter of A's being formally justified in doing \( \phi \). Formal justification, however, does not turn on real rights and genuine wrongs, as Fontaine supposes justification defenses do.

On cases falling under (1)(b), Baron shows that Fontaine's supporting case does not support him at all. I think, however, that she gives up too easily when she goes on to concede the basic point that unreasonable provocation cannot be argued as a justification. She assumes that reasonable belief and conduct is required for a good claim of justification, and this is simply not true. No more than an honest belief is often thought to be sufficient for justification—most notably in the Model Penal Code, which requires only that the defendant "believes," not that he "reasonably believes," in the justifying facts.\(^{15}\) This is a deeply entrenched feature of the Code. It generally requires subjective fault—purpose, knowledge, or recklessness—to be shown as to each material element of a crime. Criminal negligence—extreme unreasonablelessness—is disfavored and marginalized in a number of ways. This commitment carries over to the Code's justification provisions. The prosecution cannot prevail merely by showing negligence regarding non-justification; it must show at least recklessness regarding non-justification. By implication, an honest belief in justifying facts will suffice to acquit because an honest belief, even if unreasonable,

\(^{15}\) See Anthony M. Dillof, Unraveling Unknowing Justification, 77 Notre Dame L. Rev. 1547, 1555 (2002).
negates purpose, knowledge, or recklessness regarding the absence of justifying facts. This is a structure of justification so deeply embedded in the overall structure of criminal wrongdoing that it is more accurate to say it produces genuine justification, and not merely an excuse.14

Fontaine’s argument on cases falling under (2)(a) is that if D is provoked by P but D kills V, then D is guilty of manslaughter, not murder, for the death of V. He claims that in the same situation D cannot assert a justification as a defense to his killing V. But neither the case that Fontaine cites nor the case law generally says so. On the contrary, as Baron shows, the doctrine of transferred justification is perfectly plausible and widely recognized. If D is justified in defending himself from S but kills V instead, D is not guilty of the murder or manslaughter of V. The justification transfers, so to speak. Baron’s conceptual analysis is particularly enlightening here. In a case of self defense, D is formally justified and this is sufficient to acquit him regardless of material justification—that is, regardless of whether he is justified in killing the actual victim. As a result, it is a simple matter to frame the provocation argument available in cases falling into Fontaine’s category (2)(a) in terms of justification.

Baron concedes Fontaine’s point in cases falling under (2)(b). Fontaine cites a case in which the court approves the idea that heat of passion can transfer from the actual provoker to a third party. So if P provokes D and D deliberately kills V instead of P, then D is entitled to an instruction on what is best characterized as a defense of transferred heat of passion. Fontaine claims, and Baron agrees, that third-party killings such as these cannot be justified. I would add only that to make an analogy to transferred intent seems an unnecessarily awkward way of saying that the victim’s identity is irrelevant in provocation because the defendant’s diminished rational capacity holds generally for his actions in these circumstances. The best description of this case is clearly as an excuse, as Fontaine and Baron agree.

As it happens, however, Fontaine has an even stronger case to make here. The idea of transferred justification on the model of transferred intent is questionable, because offenses and justifications are deeply different. This strengthens his argument with respect to cases falling under (2)(b) and also (2)(a). The best way to understand transferred intent is to throw out altogether the ludicrous idea that intentions transfer from person to

---

person (along with the practice of arguing from metaphor gener-
ally). If D intends to kill L, then he has the requisite fault for
murder: a purpose to cause the death of a human being. If he kills
V instead, then he has caused the death of a human being. The
elements of murder are shown and D is guilty of murder. Period.
The typical statute requires only the death of and a purpose to
cause the death of “a human being.” Any dead person will do—
provided that D has caused that person’s death. In other words, so-
called transferred intent is a causation issue, specifically one of
proximate causation.

One could argue on behalf of provoked actors who intentionally
kill a non-provoking third party, and where so-called “transferred
justification” might be invoked, that the analysis is the same. If D is
justified in killing L but kills V, then he is innocent. He reasonably
or honestly believed that L presented a mortal threat, and he is
justified in causing the death of a human being. Period. Whether
or not he is guilty of the death of V is purely a question of causa-
tion—which in justification cases is moot. If he was not the cause
he is innocent by virtue of that; and if he was the cause he is inno-
cent by reason of his justification. Therefore, a case of heat of
passion in the deliberate killing of a person who has not provoked
the defendant can be seen as a justified killing.

This argument seems like a reductio ad absurdum, and it is. But
notice that what is reduced to an absurdity is the idea of trans-
ferred justification itself. This undermines Baron’s case for treating
provoked deliberate and accidental killings of third parties as cases
of justification. Fontaine prevails again, not only in cases under
(2)(b), but also in cases under (2)(a).

V. THE PROPER SEQUENCE OF EXCUSES

As Peter Westen’s title indicates, he thinks that Fontaine has
gone about arguing a point that they agree on—that provocation
cannot be a justification—the wrong way. As I see it, Westen goes
about things the right way by making better use of a wrong-headed
view of justification defenses. That is, Fontaine assumes that justifi-
cation always turns on true rights and real wrongs, omitting
mistakes. Instead of faulting Fontaine for using this conception of
justification, as Marcia Baron and Stephen Morse do, Westen
pushes it farther than Fontaine does and produces a better argument.15

Westen describes five cases that extend “classic provocation” and that can be argued only as excuses. These are lexically ordered excuse rules, each adding some feature(s) to create a more complex or more finely articulated rule that is still identifiable as an excuse. Each variation is preferable to the one before it because its greater complexity allows it to treat more cases that are plausibly viewed as excuses as representing genuine excuses. According to Westen, Fontaine purports to show that the fourth variation of provocation—the intentional killing of a non-provoking third party—conceptualizes provocation as an excuse in a way that is jurisprudentially preferable to the second and third variations because it explains how provocation reaches the reactive cases. However, Fontaine defends his preference for the fourth variation of provocation as an excuse in a roundabout fashion by defending the proposition that “provocation is an excuse.” The trouble with this argument is that the second and third versions are also excuses, and yet they are not preferable from Fontaine’s point of view. Fontaine either has not made an argument or, if he has, he has also undermined it.

Westen then makes a better argument, using Fontaine’s assumption that justification defenses always involve actual rights and wrongs. The gist of Westen’s argument is that, as the versions of excuse become ever more jurisprudentially adequate, the addition of a justification component can never enter into the process. What, asks Westen, are the implications of framing provocation as a justification on a deeds view of justification defenses? A deeds view of justification defenses takes justifying facts to be necessary and sufficient to justify, implying that one can be justified without being aware of the justifying facts. This means that justification can never be added to one of the extensions of classic provocation in a way that would make it more jurisprudentially sound, because one simply cannot act with provocation and yet be unaware of the provoking fact. A justificatory component would nullify any version of provocation with such a component if the justification argument recognizes justifying facts as necessary and sufficient to justify. Put another way, on a deeds view of justification defenses, provocation arguments exclude justification arguments.

From a slightly different vantage point, Westen’s argument is that justification is conceptually incompatible with provocation be-

cause justification excludes knowledge of justifying facts from any role whatsoever in its showing that there has been no criminal wrongdoing. Provocation gives knowledge of provoking acts a necessary, indeed central, role in mitigating criminal wrongdoing. The strength of this argument can be more fully appreciated by contrasting it with the extensional fallacy. In a pure deeds jurisdiction, there will still be cases of justification in which there is awareness of the justifying facts, among which might be provoking acts. But so what? This awareness is neither necessary nor sufficient to acquit. A pure deeds conception of justification excludes awareness of justifying facts from the argument for justification entirely, which means that awareness of provoking acts is also excluded from the argument for justification. This is certainly not true of the argument for provocation. Westen says, correctly, that on a pure deeds conception of justification, provocation cannot be a justification. Justification and provocation as arguments are mutually exclusive.

There is a response to Westen’s argument that Fontaine might make here, although the cost to Fontaine’s own argument might be too high. Notice that Westen relies not only on a deeds view of justification defenses, but on a particularly radical one. The incompatible conception of justification defenses that Westen has in mind is a pure deeds view of justification under which justifying facts are necessary and sufficient to acquit. This is a necessary assumption in Westen’s argument because if beliefs are given any role in justification defenses, then provocation arguments do not exclude justification arguments. As Pillsbury’s analysis suggests, exculpating beliefs might be given a role in provocation arguments that, with regard to our normative choices about individual responsibility, is the same role we give them in justification arguments. For example, we might treat exculpating beliefs as dispositive in both doctrines, such that beliefs alone, regardless of the actual circumstances, can suffice to establish each defense—a position on exculpating beliefs that the Model Penal Code comes very near to adopting.¹ Six On the MPC view, provocation arguments certainly do not exclude justification defenses.

Only a pure deeds conception of justification defenses, in which exonerating beliefs play no role, will guarantee that provocation arguments exclude justification arguments. The problem is that, as

far as I am aware, no one has proposed or defended a pure deeds view of justification defenses in theory or doctrine. The viable candidates for the correct conception of justification defenses suppose that awareness of justifying facts is either necessary or sufficient, or necessary and sufficient, to justify. No one seems to have argued that belief in justifying circumstances is neither necessary nor sufficient for genuine justification. It is not hard to see why. Justification defenses are to some extent continuous with or the converse of offenses because they mutually define the rights and wrongs of the matter. A pure deeds view of justification would be the counterpart of strict liability in all offenses. Just as no one has proposed that all offenses are or ought to be purely strict liability, in the sense of criminal liability’s being completely independent of criminal fault, no one has proposed the even less likely regime of pure deeds justification that acquits independently of beliefs. This is not a knockdown argument against Westen. If a pure deeds conception of justification defenses is correct, then his argument holds up. But this is, I think, an awfully big assumption. Unlike Westen, Fontaine might be able to use a negative elements analysis of justification defenses, because on that view all justification defenses at least can be established by facts alone, which is not true of provocation.

VI. Fallacy and Morality

Stephen Morse wastes no time targeting the two principal weaknesses of Fontaine’s argument. First, Fontaine claims that his account of provocation is definitive because it is a claim to analytical truth—but the arguments that purport to establish this analytical truth are logically invalid. Second, Fontaine’s argument against reading provocation as a justification is that justification

18. On a negative elements view of justification defenses, one who is mistaken in his beliefs about justifying facts is no more guilty than if his mistake of fact pertained to the elements of the offense, and for the same reasons. In the proof of non-justification, the prosecution has to prove pairs of fault and non-justification elements, implying that the defendant prevails if she can prove either a fact element of a justification defense or non-fault (i.e., a reasonable or honest belief) regarding such an element. As is true in the proof of offenses, justifying beliefs are sufficient to acquit—as are justifying facts, regardless of beliefs. See Huigens, supra note 14.
19. MODEL PENAL CODE § 3.02(1) (1962).
supposes the truth of its factual bases, which means that cases of adequate non-provocation cannot be accommodated under the justification umbrella. The problem, Morse points out, is that these cases can indeed be analyzed as exhibiting genuine justification, and not merely an excuse premised on mistake—which means that there is no obstacle to formulating provocation arguments as justification arguments. Morse's main point—that if provocation is an excuse and not a justification, then this reflects a normative choice, not an analytical truth—follows from these observations.

Morse takes away Fontaine's premise that justification requires vindicating true rights and real wrongs. This is not true if mistaken justification can be genuine justification, and Morse makes a straightforward moral argument that it can be. One who is mistaken about justification is genuinely justified because "[a]n agent who acts entirely reasonably for a justificatory reason has done all that any decent society can expect of him and therefore he is justified despite the regrettable mistake." It is worth considering a more detailed version of this argument that Morse has published previously:

If the agent's reasonable belief that justification obtains is discovered ex post to be mistaken, then that is a cause for great regret, but not for condemnation [or] guilt .... The reasonably mistaken person behaved flawlessly [and] there has been no wrongdoing. The agent caused unnecessary and regrettable harm, but this does not indicate that the agent violated any moral or legal expectation. Indeed, under similar circumstances, the law would ex ante encourage her to do it again. A reasonably formed belief does and should give the agent good reason to act, because reasonably formed beliefs are likely to be true according to our best standards of evaluation and truth.22

It seems to me that Morse has by far the better argument here. As he points out, however, the mere fact that this view is widely held is sufficient to disprove one of Fontaine's critical premises.23 As Morse notes, it also casts doubt on Fontaine's claim to have established that provocation is an excuse as an analytical matter.

21. Id. at 200.
In a couple of respects, however, Morse's argument is weaker than it needs to be. To describe reasonable mistakes regarding justification as causing "unnecessary and regrettable harm" can be read to describe the mistake as an excuse. Morse assumes that doing harm can be distinguished from wrongdoing, so that to make a mistake about harm done in justification nevertheless results in a genuine justification. But it is possible to argue that harm-doing is not so easily distinguished from wrongdoing. Roughly speaking, a consequentialist theory of punishment will assert that criminal wrongdoing does not extend beyond harm-doing. If doing harm and doing wrong are not distinguishable, then it is open to Fontaine to argue that to read his mistake in provocation cases as mistake in justification cases is nevertheless to read the former cases as excused. This does not prove his argument, and it does not rebut Morse's point that the mere existence of alternative views casts doubt on his claim to have established an analytical truth. It does, however, give Fontaine the prospect of calling in allies: a group consisting of George Fletcher,25 Heidi Hurd,26 Russell Christopher,27 Peter Westen,28 and John Gardner29 among many others who take at least some mistakes in justification to result in an excuse. As Westen himself shows in his contribution to this Symposium, these are valuable allies, to say the least.

Fontaine has another answer to Morse, one that is worth noting though it, too, is not a knockdown argument. In his essay for this Symposium, as in his earlier analysis of the question, Morse consistently refers to reasonable mistakes, implying that an unreasonable mistake results in an excuse, not a genuine justification. Fontaine might argue that Morse's conception of justification defenses covers only two of his four cases: reasonable mistakes about provoking circumstances and accidental killings in response to provocation. This leaves Fontaine's argument intact where his other two cases of actual non-justification are concerned. If unreasonable mistake and the deliberate killing of third parties are merely excuses, then these cases, at least, imply that provocation cannot be a justification. The problem with this argument on Fontaine's behalf is that the Model Penal Code differs from Morse on this point. Only an

24. Morse, supra note 22, at 407.
27. Christopher, supra note 17, at 251.
honest belief, not a reasonable belief, is required for genuine justification under the Code. The Code, at least, rebuts Fontaine's premise for all four of his actual non-justification cases, because it treats both reasonable and unreasonable mistakes as genuine justifications. Again, as Morse points out, the mere existence of this alternative view casts doubt on Fontaine's claim to having uncovered an analytical truth.

Regarding Fontaine's flawed logic, Morse focuses on the following passage: "[I]f the doctrine were one of justification . . . , it would necessarily mean that these courts, the corresponding statutory law, and the MPC all have heat of passion dead wrong, as heat of passion can only be argued to be a partial justification . . . where there exists serious provocation . . . ."30

Morse responds:

If provocation/passion is a partial justification, then it is a claim that a killing in these circumstances is less wrong and more acceptable than if there were no or inadequate provocation deserving of a return . . . Denying this possibility simply begs the question against the partial justification version and does not analytically demonstrate that the doctrine must be a partial excuse.31

Morse is correct that Fontaine's argument begs the question, but it is worth considering how, exactly, Fontaine goes wrong. It may be that the damage is repairable.

Putting the passage that Morse cites in its larger context, it is evident that Fontaine begins with a perfectly good argument that brings him to the flawed one. The good argument goes like this:

If courts are not wrong to call heat of passion an excuse and apply it to reactive-homicide scenarios, then heat of passion is an excuse.

Courts are not wrong to call heat of passion an excuse and to apply it to reactive-homicide scenarios.

Heat of passion is an excuse.

This is a valid argument of the form modus ponens \( p \supset q, p, \therefore q \), even if it is not very informative. But Fontaine goes wrong at the next

30. Fontaine, supra note 1, at 42 (emphasis added).
31. Morse, supra note 20, at 202.
step. If his argument is to be informative, then he needs to obtain the conditional premise for this argument—the proposition that courts are not wrong to call heat of passion an excuse and to apply it in reactive scenarios—and he argues for this as the conclusion of the argument that Morse quotes.

Before I analyze that argument, let me make two preliminary points. First, I take the gist of Fontaine’s argument to be that justification arguments can be made only where genuine justification appears. Therefore, justification arguments cannot be made in cases of reactive emotions and provocation interpretational bias (PIB). It follows that heat of passion cannot be a justification in reactive/PIB scenarios. So for clarity’s sake, I will clean up one of Fontaine’s propositions as follows: Heat of passion can only be argued to be a justification where there exists a serious provocation \( \equiv \) Heat of passion cannot be a justification in reactive scenarios. Second, the word “as” in the sense of “because” in an argument must be read as a premise of the argument, and cannot be read as the argument’s conclusion. This may be obvious, but it is also important.

The argument quoted by Morse goes like this:

If heat of passion is a justification in reactive homicide scenarios, then courts are wrong to call heat of passion an excuse and apply it to reactive-homicide scenarios.

Heat of passion cannot be a justification in reactive homicide scenarios.

Courts are not wrong to call heat of passion an excuse and to apply it to reactive-homicide scenarios.

This is an invalid argument because it denies the antecedent. It says: \( p \supset q, \neg p, \therefore \neg q \). This is like saying: If Bill stole the church funds, then he would be a felon; Bill did not steal the church funds; so Bill is not a felon. This seems valid because Bill is not a felon by virtue of having committed this particular theft. But it does not follow that Bill is not a felon full stop. For all we know, he has a sheet of felony convictions as long as your arm. Likewise, Fontaine forgets about the world of possibilities outside his conclusion. It may be that courts are not wrong to call provocation an excuse and apply it in reactive scenarios. But this might be true because courts correctly analyze provocation in some way that is neither here nor there where provocation’s being a justification in reactive homicide scenarios is concerned. For example, these courts might be right about provocation’s applying in reactive
situations because they correctly treat it as an inquiry into the quality of the defendant's practical reasoning, as Gardner portrays it. This leaves open both the possibility that provocation can be interpreted as a justification in reactive homicide scenarios and the possibility that it cannot be. Gardner’s interpretation rules out only taking provocation as a claim of non-responsibility.

Returning to Fontaine’s original, valid argument, we now find that he has not supported it. It is a valid argument, but so far it just is not informative. Fontaine has not established his premise that courts are not wrong to call heat of passion an excuse and apply it in reactive scenarios. But perhaps Fontaine’s argument can be salvaged. It certainly sounds plausible to say that courts are not wrong to call heat of passion an excuse and apply it in reactive scenarios. If Fontaine goes wrong in denying the antecedent, perhaps he could simply not do that and obtain a valid argument.

If heat of passion is a justification in reactive homicide scenarios, then courts are wrong to call heat of passion an excuse and apply it to reactive-homicide scenarios.

Courts are not wrong to call heat of passion an excuse and to apply it to reactive-homicide scenarios.

Heat of passion is not a justification in reactive homicide scenarios.

This is a valid argument in the form of modus tollens: $p \supset q, \neg q, \therefore \neg p$. The problem is that this argument does not support the first valid argument that Fontaine set out to support. He needed to obtain the conclusion that courts are not wrong to call provocation an excuse and apply it in reactive scenarios. But in this second, valid argument, that proposition is only a premise.

As I said above, it might be possible to salvage Fontaine’s arguments. I think it can be done by recasting them in a way that can open up his thesis in a way that better accommodates the social scientific literature.

VII. ABDUCTION, PUNISHMENT THEORY, AND HEEDING WEISBERG

An odd thing happened as I read these essays—particularly Baron’s and Morse’s—and as I re-read (and re-read) Fontaine’s article. It began to seem that he was not engaged in the theory of punishment at all. The theory of legal punishment is
predominantly if not exclusively a branch of philosophy, drawing on moral philosophy, general jurisprudence, and, more recently, the theory of action. It seemed to me that this was not what Fontaine was doing. He seemed to be doing social science. Unconsciously, I think, he claims to have produced a definitive account of provocation as an excuse because he has produced a strong hypothesis to that effect, and because strong hypotheses are not uncommonly described by scientists as the truth of the matter. Given that scientific hypotheses are meant to be falsified, this is an odd thing. It is, nevertheless, characteristic of the scientific community.

When I say that some of Fontaine's arguments are by nature scientific, I mean something specific. Some of them are abductive arguments. These arguments are pervasive in science and perfectly sound in that context because abduction is the kind of reasoning by which hypotheses are formed—even though the same arguments are fallacious as a matter of deductive logic. Recall this example of the fallacy of denying the antecedent: If Bill stole the church funds, then he would be a felon; Bill did not steal the church funds; so Bill is not a felon. It is easy to mistake this for a valid argument because if Bill did not steal the church funds, then he is less likely to have felony convictions. It is not a deductively valid argument, but from an empirical point of view, if we want to know whether Bill is a felon, his handling of the church funds is a good place to start looking. His not stealing the church funds is a piece of evidence that he is not a felon. An abductive argument does more than to state the hypothesis. An abductive argument can be better or worse. The otherwise question-begging argument might be well framed to guide an empirical inquiry because it is sensitive to context, because it is clear how it might be falsified, and—most importantly—because it will indicate how we might interpret its falsification. If a hypothesis seems likely but is falsified,

32. For example, this practice accounts for much of the confusion over the teaching of the theory of natural selection in schools. The advocates of teaching creationism as an alternative constantly attempt to denigrate natural selection by calling it "only a theory" and equating creationism with it as an alternative theory, on the ground that neither one has been definitively proven. This is wrong because it overlooks the fact that "natural selection is a theory" and "natural selection has never been definitively proven" are entirely consistent with the claim that natural selection is true. This is simply the nature of scientific theories: they consist of hypotheses that always remain subject to falsification. Scientific truths are true only in the sense that they have not yet been falsified. Natural selection is extraordinarily resistant to falsification and creationism falls apart in a light breeze. The former might be false, but the latter is definitely false. The scientific community would do the opponents of teaching creationism a great service if they would refrain from describing natural selection as true, or at least make the nature of scientific truth clear to the lay people engaged in this political debate.
then we know we need to make radical adjustments to it or abandon it all together. If it sounds crazy initially, but we cannot falsify it, try as we might, then it is more and more sensible to treat it as fact.

The following otherwise invalid argument can be treated the same way as an abductive argument:

[Cases] reflect a longstanding rationale as to the excusatory nature of the doctrine. The rationale is that if the doctrine may be conceptualized and applied as an excuse, then, even if it is represented and can be explained, at times, as a justification, it cannot be the latter because the former "admits to the existence of a social harm" and to the wrongfulness of the act. In contrast, a justification is an assertion that the act is acceptable and makes no such concession. In this way, excuse is a lower moral standard than justification, and, as such, the defense cannot be a justification (requiring it to always meet this higher moral standard) if it can, at any time or in any instance, be rationalized and viewed as an excuse.\(^{33}\)

The argument of this paragraph is: if provocation can be an excuse, which makes a weak moral claim, then it cannot be a justification, which makes a strong moral claim; provocation cannot be a justification, which makes a strong moral claim; so provocation can be an excuse, which makes a weak moral claim. In other words, \(p \supset q, q, \therefore p\). Fontaine's argument affirms the consequent, and is deductively invalid. But it also states a good hypothesis. It is framed narrowly enough to be looked into, and is likely enough, in context, to merit looking into because it is a substantial claim for which there is good and bad evidence to be had. Do the cases on provocation consistently make only the weak moral claims characteristic of excuses? Are courts that make strong moral claims for provocation consistently overruled? If so, having failed to falsify that hypothesis, then we might arrive, by induction, at the conclusion that provocation is an excuse, which makes a weak moral claim.

To bear in mind that Fontaine is a social scientist makes sense of his tendency to make bad deductive arguments that are good abductive arguments. It also makes sense of an odd statement that he makes in passing—and one that deserves attention. He writes: "The fact that many courts have understood heat of passion as a

\(^{33}\) Fontaine, supra note 1, at 42 (quoting Joshua Dressler, Rethinking Heat of Passion: A Defense in Search of a Rationale, 73 J. CRIM. L. & CRIMINOLOGY 421, 438 (1982)).
partial excuse, and have applied the doctrine across varied reactive-homicide scenarios, does not in itself make the doctrine one of excuse (i.e., the doctrine's nature exists independent of said cases). The truth is that the doctrine's nature does not exist independently of the cases. The cases constitute the legal doctrine. Even outside common law, when courts interpret statutes and constitutional provisions, the cases are not merely evidence of what the law is. They are what the law is. The fact that Fontaine sees cases as evidence of legal doctrine is another indication that he is thinking more like a scientist and less like a legal theorist.

Fontaine might embrace the former role, and if he did then it would be perfectly fine for him to treat the cases as evidence of what the law is. The scholarship and commentary surrounding the doctrine might also be taken as evidence of its nature, as might the moral judgments that surround the doctrine—especially given that much of the scholarly commentary makes a moral case for its interpretation of provocation. For example, it has been argued that provocation should not be treated as a justification because to do so revives the doctrine as it existed when it was almost certainly seen as a justification—that is, when it endorsed honor killings and the patriarchal treatment of women. The fact that these are archaic doctrines that have been rendered obsolete, as a matter of historical fact, by changes in the wider culture is evidence that provocation is not a justification. If the cases, legal commentary, history, and informal moral judgments surrounding provocation support Fontaine's thesis in the role of evidence of provocation's nature, then perhaps he should approach provocation as a matter of sociology.

It is clear that this is not the direction in which Fontaine wants to go. To see where he does want to go and how he might get there requires an understanding of the theory of legal punishment as normative theory. Some legal theory does support one policy or doctrinal recommendation over others. H.L.A. Hart's *Law, Liberty, and Morality*—in which he made the case against the criminalization of homosexuality—is usually read this way. But in most punishment theory, its normative dimension is implicit. It is taken as a given that the correct description of legal punishment is a description of just punishment. It is far from obvious, however, how this works. I think that most punishment theorists make two

34. *Id. at* 41–42.
assumptions that they are not aware of, for the most part, and that they might reject if these assumptions were brought to their attention. They suppose that just punishment supervenes on legal punishment, so that to describe the latter correctly is to describe the former.\textsuperscript{36} They also adhere to a weak internalism, in that they suppose that to take these descriptions of law and justice to be true is to have reasons to make them so.\textsuperscript{37} On these assumptions, justice requires us to adopt law as the theory describes it as positive law. If the theory of legal punishment is "normative theory," then it is usually normative in this sense, and not in the simple sense represented by Law, Liberty, and Morality.

How does this help Fontaine? Note that there is a major disconnect between his article and the commentary in this Symposium. He brings the social science of provocation interpretational bias to bear on the cases and attempts to use it to support his interpretation of provocation. With one exception, the participants in the Symposium marginalize this aspect of his argument. On the other hand, I think that Fontaine has not used the science as effectively as he might have done. The key to strengthening his argument, I think, is to bear in mind the intersection of description and normativity that is implicit in the notion of normative legal theory. He attempts to find the points at which the social science and the law match or make contact. Hence his scientific approach to both, including his assumption that the cases are evidence of the law. He ought, instead, to think of the goal as being to show that the phenomena that the social science describes are to be found in the law. This is an interpretive approach instead of an empirical one, and it is characteristic of the theory of legal punishment. If the law is as Fontaine describes it, and if the cases constitute instead of evince the law, then there is, in at least some provocation cases, an implicit recognition of reactive emotion and provocation interpretational bias; an implicit recognition that such cases cannot support a justification argument; and so on. As a theorist of punishment, Fontaine's job is to tease this recognition out of the cases. This is different from treating the cases as confirming instances of his thesis. That he has not approached the theory of punishment as the commentators expected him to might explain some of the disconnect.


This view of things helps Fontaine because it shows, at least, that he is not wrong. Most of the cases he examines are, in fact, confirming instances of his thesis. But to take a theoretical approach, to think of the cases interpretively, as instantiating his thesis, will significantly widen the cases that are relevant to it, and that might support it. More important, his interpretation of provocation as excuse plainly has a normative dimension to it. Shifting his approach from a scientific one to the interpretive project that is more characteristic of punishment theory gives him the advantage of the distinctive way in which description and normativity intersect in the normative theory of punishment.

Fontaine has the great good fortune to have Robert Weisberg’s essay to point him in the right direction. I believe that Weisberg describes, in a slightly different way, the difficulties in Fontaine’s approach that I have described here. Weisberg takes Fontaine’s enterprise to be a scientific one, and in one aspect in particular Weisberg urges caution in carrying it out. This has to do with the “mapping” of hypotheses onto evidence.

In *Disentangling the Psychology and Law of Instrumental and Reactive Subtypes of Aggression*, Professor Fontaine is quite anxious to draw on psychological resources. He seems tempted to rely on the established reactive/instrumental distinction because it might support his legal argument for (a) treating HOP as an excuse and (b) including PIB cases under the voluntary manslaughter umbrella. On the other hand, he is at least skeptical about the validity of the distinction (if more tentative in his view than are Bushman and Anderson), noting that there are overlaps in the incidences and developmental etiologies of the two forms of aggression. Indeed, Professor Fontaine refers to quite a number of “dichotomous models of aggression” to complete the picture of the current scientific research. His discussion of these is rich and nuanced. But even while he is properly cautious about expecting scientific distinctions to map easily on to legal ones, he does, like Bushman and Anderson, tend to analogize the key dichotomies (involving, whatever the vocabulary, passion and emotion) to the legal distinction between “malice aforethought” and “heat of passion” and thereby assumes away the problem of providing some affirmative conception of second-degree murder. The temptation he shares with them to view

this as the salient legal distinction suggests that he still struggles with whether and how to address the mapping difficulty.\textsuperscript{39}

It is important to note that Weisberg views the state of Fontaine’s body of work in a more positive light than I have done, just as he views the most recent installment far more positively than do the other contributors to the Symposium. For what it is worth, I think that Weisberg captures the scientific orientation, spirit, and value of Fontaine’s article.

At the same time, Weisberg seems to make the same recommendation that I do—that is, that Fontaine needs to better integrate his work with punishment theory as others practice it. In the course of describing Peter Arenella’s work on law and moral luck, Weisberg writes:

\begin{quote}
[Arenella] recognizes the fundamental “control principle,” i.e., that people should not be held responsible for events or circumstances they cannot control, and he laments that many philosophers adhering to the control principle draw an arbitrary line by failing to acknowledge that some offenders do not have control-act freedom over how to do and how to be. Arenella recognizes that he is thrusting himself into knotty debates over “hard” vs. “soft” determinism. He would not excuse the psychopath for whom moral judgment and moral motivation are simply much more difficult than for others; he worries about those for whom, as a threshold matter once adult legal responsibility sets in, they are impossible. So if Professor Fontaine is committed to a coherent psychological explanation of PIB and to giving it legal valence, I would press him to consider how it compares to Arenella’s depiction of the moral desert of the amoral psychopath.\textsuperscript{40}
\end{quote}

As Weisberg describes it, Arenella consciously worked against the mapping of philosophical conclusions onto the law and theory of punishment as one would map hypotheses onto evidence. Instead, Arenella sought to inject empirical evidence into the moral philosophical debate in the true spirit of that debate. Weisberg plainly recognizes the difficulty of doing this, as do I. Weisberg seems confident that Fontaine is up to this difficult task, and I share that confidence. Weisberg obviously is enthusiastic about Fontaine’s

\textsuperscript{39} Id. at 64–65 (quoting Reid G. Fontaine, Disentangling the Psychology and Law of Instrumental and Reactive Subtypes of Aggression, 13 PSYCHOL. PUB. POLY & L 143, 153–55 (2007)).

\textsuperscript{40} Id. at 73–74.
project and looks forward to the next installment. I will venture to say that all the participants in this Symposium share those feelings.