Reactive Cognition, Reactive Emotion: Toward a More Psychologically-Informed Understanding of Reactive Homicide

Reid G. Fontaine, University of Arizona
REACTIVE COGNITION, REACTIVE EMOTION
Toward a More Psychologically-Informed Understanding of Reactive Homicide

Reid Griffith Fontaine
University of Arizona

Recent scholarship has drawn attention to the alternative contributions of dysfunctional reactive cognition (e.g., provocation interpretational bias) and emotion (e.g., provoked fury) in heat-of-passion killings. Two main theses have been advanced. First, there exists a meaningful parallel between the instrumental–reactive aggression dichotomy in psychology and murder–manslaughter distinction in law. Second, analysis of this parallel suggests that the heat-of-passion (or provocation) defense disproportionately favors emotional over cognitive dysfunction in mitigating murder to manslaughter. These theses, however, have yet to be fully developed and raise additional, critical questions that have not yet been addressed: Other than interpretational style, how may social–cognitive science inform our understanding of the role of cognitive bias in reactive homicide? How is serious interpretational bias related to alternative forms of psychiatric disorder as recognized in law? This article addresses these and related questions regarding the differential and interactive contributions of dysfunctional cognition and emotion in the execution of reactive homicide.

Keywords: instrumental–reactive aggression, heat of passion, murder–manslaughter distinction, provocation, cognition

Recent scholarship has drawn a parallel between the instrumental–reactive aggression dichotomy in psychology and the murder–manslaughter distinction in law.1 Whereas instrumental aggression is characterized as cold-blooded, nonemotional, and premeditated for the purpose of personal gain, reactive aggression is typically hot blooded, emotionally charged, and enacted quickly for the purpose of harming a perceived provocateur or defending oneself. The instrumental aggressive subtype represents malice aforethought murder in criminal law, by which the killer premeditates the crime against his or her victim. Alternatively, the reactive aggressive subtype reflects heat-of-passion manslaughter, by which the killer takes his or her victim’s life in a state of reactive fury, a state that is triggered by his or her belief that he or she was first provoked by the victim.

This parallel has raised important scholarly questions about the nature of provoked killings and how legal doctrine may and should be informed by psychological science. In particular, research on the relation between biased

The Heat-of-Passion (or Provocation) Defense

The practice of excusing a killer because the homicide was committed under conditions that are now recognized by the heat-of-passion doctrine existed long before the doctrine was officially established. Jeremy Horder provided a detailed history of the origins of *provocation law* (the justice and punishment of alleged crimes that were reported to have been committed in response to some form of provocation) and depicted the medieval jury as a select group of local men who were believed to have knowledge of the defendant and the alleged crime. In cases in which the defendant was provoked but the requirements of self-defense were not met, juries would often embellish the facts to make them fit the excuse of self-defense; that is, in the case of the defendant who killed in reaction to a provocation but did not need to kill to save his or her own life (or the life of another), it was not unusual for the jury to rearrange or adjust the circumstances of the killing to interpret the situation so that the defendant killed in self-defense. This is remarkable in that, in these cases, juries felt so strongly that a defendant should not be charged with felonious homicide (i.e., murder) when the killing occurred out of provoked rage that they creatively reworked the history of the incident so that the defendant’s crime may be entirely excused.

Subsequently, it became better recognized that provocation killings did not properly fit into either the category of murder or the category of self-defense. It was understood that, whereas provoked killings were not committed in cold blood with the malice aforethought that was requisite of murder, they were indeed acts that were more morally reprehensible, and likewise more criminally culpable, than homicides that were committed in self-defense. As a result, the heat-of-passion defense was born.

In American common law, the heat-of-passion defense reduces a charge of murder to manslaughter if the defendant successfully demonstrates that he or she (a) was first adequately provoked by his or her victim, (b) became emotionally disturbed

---


4 Note that modern criminal law has viewed killing in self-defense as a justifiable, as distinguished from excusable, homicide.


6 This article focuses on the traditional American common law framing of heat of passion. The applicability of the arguments set forth herein to heat of passion/provocation in other Anglo-American jurisdictions varies. Also, related issues that are specifically relevant to the Model Penal Code’s “extreme mental or emotional disturbance” language [Model Penal Code § 210.3(1)(b) (1962)] have been addressed elsewhere. See Reid Griffith Fontaine, *The Wrongfulness of Wrongly Interpreting Wrongfulness: Provocation Interpretational Bias and Heat of Passion Homicide*, New Crim. L. Rev. (in press).
as a direct result of said adequate provocation, (c) did not have enough time to “cool
off” before killing his or her perceived provocateur, and (d) did not, in fact, cool off
before committing the homicide in question.\(^7\) Noteworthy is that, although the
criterion of *adequate provocation* needs to pass the “objective” reasonable-person
standard (i.e., a reasonable person, or person exhibiting understandable judgment
given the circumstances, would have interpreted the social stimulus in question to be
sufficiently provocative to cause one to suffer emotional disturbance), the requirement
that the defendant experience emotional disturbance need only pass the less stringent
subjective standard. In psychological terms, the defendant’s cognitive functioning (or
interpretation of the stimulus/provocation) is held to a more strenuous standard than
is his or her emotional functioning (or reactive fury). This raises both logical and
moral questions about the structure of the heat-of-passion doctrine and creates a
disparity such that, everything else equal, the heat-of-passion killer who has signifi-
cant cognitive difficulties (particularly those of an interpretational nature) is ascribed
greater culpability and more severely punished than the cognitively intact heat-of-
passion killer. On the basis of the current framing of the defense, the former’s
invocation of the heat-of-passion defense fails, whereas the latter’s succeeds.

Consider the following example of a situation that is open to being interpreted
as a provocation: A husband learns that his male friend has been visiting his wife
at their home while he is at work. Unbeknownst to the husband, his friend’s
purpose has been to help his wife plan his surprise birthday party. The husband
returns home to talk with his wife and inquire about the nature of the situation.
However, upon reaching his property, he discovers his friend entering his home.
At the sight of this, the husband becomes enraged and, in an act of fury, storms
into his house and kills his friend.

In this scenario, the friend’s behavior is open to multiple interpretations, one
of which is that the friend is having an affair with the husband’s wife. However,
it is likely that the husband’s belief would not pass the reasonable-person
standard. Rather, the reasonable person would likely consider alternative expla-
nations as to the friend’s visit or at least think to first query his friend and wife
before acting, which, in turn, may prevent a heated reaction of violence. If the
husband, however, suffered a history of betrayal by loved ones, he may be
strongly biased toward quickly and definitively concluding that his friend and
wife are betraying him. Cognitive dysfunction, such as the bias here illustrates, is
not accommodated by the common-law heat-of-passion doctrine.

The difficulty in considering cognitive bias in reactive homicide is illustrated by
the recent Alaska case, *Howell v. State*.\(^8\) In this case, Howell, the defendant, reported
that he had been shot at earlier in the day at the close of an interaction with another
male adult. Later that evening, in a separate series of social exchanges, the defendant
took to intermittent arguing with a different male acquaintance named Church, who
had visited Howell and his girlfriend at their home. At the close of their argumenta-
tion, Howell told Church what had happened to him earlier in the day and that another
man had shot at him. Church told Howell that he did not believe his story, to which
Howell responded with great displeasure. Still, the men decided that their argument

\(^7\) See WAYNE R. LAFAVE & AUSTIN W. SCOTT, CRIMINAL LAW (2d ed., 1986).

was finished, and Church left Howell’s home. After a period of time passed, however, Church returned to urge Howell, “Just forget it ever happened, just forget the whole thing.” Perhaps because he was primed to interpret the event in such a way because of having been assaulted earlier in the day, Howell interpreted Church’s communication to mean that Church was a friend of Howell’s assailant and, as such, that Church was trying to protect the assailant. This belief, reasonable or not, caused Howell to become furious (and perhaps fearful), and he shot and killed Church as the latter was leaving the area. The trial court judge, having himself decided that Church’s actions did not qualify as adequate provocation, refused to instruct the jury on heat of passion, and Howell was convicted of first-degree murder.

Is an interpretation of serious (or, in doctrinal terms, adequate) provocation, such as that made in Howell, reasonable? This, of course, is a matter to be determined by the fact finder. However, interpretations of this sort, especially when shown to be objectively incorrect, may have difficulty passing the reasonable-person standard. In the case that the interpretation is deemed unreasonable (or, more accurately, nonreasonable), the question arises as to whether the nonreasonableness of the defendant’s interpretation results from individual functioning and capacities for which he or she can be rightly held responsible (such as in the case of nonculpable cognitive dysfunction).

It is unclear, and, indeed, remains a matter of debate, as to why, in this unique instance, the law disproportionately favors emotional over cognitive dysfunction. One explanation, however, is that if this distinction between cognitive and emotional dysfunction did not exist, it would open the defense up to a larger number of bogus invocations. It may be argued that, in addition to a defendant being able to embellish extreme emotional disturbance, a partial allowance for cognitive dysfunction in addition to extreme emotional disturbance would allow the defendant greater leeway to embellish mental dysfunction in that he or she could now attempt to demonstrate that, regardless of the objective reality of the situation, he or she was unable to interpret the situation as anything other than a substantial provocation. This rationale, however, is irrelevant to the retributive basis of heat-of-passion doctrine, as it neglects the principle of penal proportionality. Presumably, the underlying principal purpose of the heat-of-passion doctrine is to partially excuse killers who have committed reactive homicides without the malice aforethought or cold-blooded premeditation (i.e., a more serious and wicked mens rea) that is elemental to murder. If so, the distinction that is inherent in the heat-of-passion doctrine between emotional and cognitive dysfunction is both logically inconsistent and morally problematic.

---

9 Id.
10 Id. On appeal, it was determined that the trial court erred in its failure to instruct the jury on heat of passion, and that the reasonableness of the defendant’s belief of provocation (i.e., whether there was adequate provocation) is a question to be answered by the fact finder. The conviction was vacated, and the case remanded.
11 By nonculpable cognitive dysfunction, I mean cognitive dysfunction that is (a) not the creation of the actor’s own doing, (b) outside the actor’s control, and (c) either not known to the actor or is known to the actor but could not have reasonably been foreseen by the actor to play a significant role in the reactive killing in question.
12 For a recent analysis and discussion of the nature and structure of heat passion, see Reid Griffith Fontaine, Adequate (Non)Provocation and Heat of Passion as Excuse Not Justification, U. Mich. J.L. Reform (in press).
13 I do not want to overstate the implications of the arguments herein set forth. It is unlikely
Reactive Cognition, Reactive Aggression

There is considerable empirical evidence that certain kinds of cognitive limitations play a critical role in reactive violence in humans. To the degree that a person is cognitively limited and biased in favor of interpreting provocations in ambiguous situations, it may not be reasonably said that he or she has a moral duty to consider alternative, nonprovocative explanations of said situations. In this way, reactive killings that are predicated on cognitive dysfunction that favors interpretations of provocation may be viewed as less criminally culpable.

A substantial amount of scientific support for the hypothesis that cognitive (or, more specifically, interpretational) bias underlies reactive anger and violence stems from research on the role of social information processing (SIP) in the development of antisocial behavior. Social–information processing theory states that social behavior is the result of multiple series of mental processes that are activated—either consciously or otherwise—in real time (or “in the moment”) to make meaning out of social cues and stimuli as they enter one’s perceptual sphere. Nicki R. Crick and Kenneth A. Dodge have articulated five steps (or domains) of cognitive operating in their reformulated model of SIP: First, the responding individual perceives and organizes various aspects and features of the stimulus (called encoding of cues). Second, the respondent interprets the meaning of the social cue and makes attributions as to causality and the stimulus actor’s intent, social style, and affect (interpretation of cues). Third, the respondent may identify preferred outcomes of the situation (clarification of goals). Fourth, alternative ways of responding to the stimulus may be generated or drawn from memory (response access or construction). Finally, response options may be evaluated across multiple domains so that the respondent may select a behavior for enactment (response decision).

As I have discussed in detail elsewhere, the first two steps of SIP—encoding and interpretation of cues—have been most closely linked with reactive aggressive behavior. Empirical studies have demonstrated that reactive aggressive

---

14 For recent reviews and discussions, see Reid Griffith Fontaine, Evaluative Behavioral Judgments and Instrumental Antisocial Behaviors in Children and Adolescents. CLINICAL PSYCHOL. REV., 956 (2006); Reid Griffith Fontaine, Disentangling the Psychology and Law of Instrumental and Reactive Subtypes of Aggression, 13 PSYCHOL., PUBLIC POLICY, AND LAW, 143 (2007).
15 This is true, of course, only to the degree that the person is not responsible for being so cognitively limited.
17 Id.
18 Id.
19 Id.
20 Id.
21 Id.
individuals are more likely to both encode (or perceive) negative aspects of social cues and attribute hostile meaning to incoming social information in ways that promote aggressive reactivity. For example, individuals who are biased in favor of perceiving neutral faces as angry and perceiving ambiguous provocation situations as intentionally hostile and harmful are more likely to demonstrate patterns of reactive aggressive behavior. This combination of cognitive processing difficulties has been called provocation interpretational bias (or PIB).23

In its most serious form, PIB can mean that the reactor is unable (or at least substantially less able, compared with others) to interpret ambiguous provocation situations as anything other than substantially provocative—that is, the defendant who has committed a reactive homicide may have been unable to not make a mistake of fact as to the provocative nature of the social stimulus, even though his or her mistaken interpretation of provocation may indeed be deemed unreasonable (and thus fail the reasonable-person standard). To briefly return to the previous example, the husband who kills his friend after watching the latter enter his house may immediately and definitively conclude that his friend is having adulterous relations with his wife because of his serious cognitive dysfunction, or PIB.

Numerous studies have investigated the relation between biased attributional and interpretational processing and antisocial behavior.24 Although much of this research has been conducted with normative samples of children and adolescents, there is no reason to expect that the relation is any less strong or less relevant to aggressive reactivity in adults. Recent research with adult samples is consistent

23 Reid Griffith Fontaine, The Wrongfulness of Wrongly Interpreting Wrongfulness: Provocation Interpretational Bias and Heat of Passion Homicide, New Crim. L. Rev. (in press). Note that the literature in psychology more commonly references the term “hostile attributional bias” (or “hostile attributional style”), which reflects aggressive individuals’ proneness to attribute hostile intent to social stimuli in ambiguous provocation situations. I use the term provocation interpretational bias because it is more inclusive and better reflects both the (a) combination of encoding-of-cues and interpretation-of-cues steps of social information processing, and (b) multicomponental nature of biased cognitive processing that aggressive individuals tend to demonstrate when faced with social situations that are ambiguous as to the provocative nature of their content. To be clear, though, a notable portion of the empirical research that supports the term provocation interpretational bias has specifically examined hostile attributional processes. Indeed, the terms provocation interpretational bias and hostile attributional bias are consistent with each other. The critical difference is that the term provocation interpretational bias consumes the processes represented by the term hostile attributional bias, and, in addition, reflects perceptual and encoding difficulties (captured by step 1 of SIP, encoding of cues) that have also been linked to individual differences in human aggressive behavior [e.g., see Kenneth A. Dodge & Joseph P. Newman, Biased Decision-Making Processes in Aggressive Boys, 90 J. Abnormal Psychol. 375, 375–379 (1981); Karen R. Gouze, Attention and Social Problem Solving As Correlates of Aggression in Preschool Males, 15 J. Abnormal Child Psychol. 181, 181–197 (1987)]. A compelling argument that emphasized the potential importance of encoding of cues to understanding individual differences in hostile attributional processes and antisocial behavior was made recently by Dr. Tako Horsley [see Tako Horsley, Menno van der Schoot, & Bram Orobio de Castro, An Eye-Tracking Study of Encoding in Aggressive Children: Does a Bias in Encoding Affect Interpretation or Is It Vice Versa? Paper presented at the 20th biennial meeting of the International Society for the Study of Behavioural Development in Würzburg, Germany (2008)].

with the youth literature. For example, scientists have recently explored hostile attributional bias in adult male offenders.\textsuperscript{25} Their findings suggest that hostile attributional bias may be associated with two maladaptive pathways—psychopathy and depressive functioning—and provide support for the hypothesis that aspects of PIB are associated with actual antisocial syndromes.\textsuperscript{26}

Still, there is considerable research in social–cognitive science that bears relevance for the link between cognitive bias and heat-of-passion homicide that has yet to be fairly addressed. For instance, L. Rowell Huesmann and his colleagues contributed to research on social cognition and antisocial behavior by focusing on the role of hostile-world schemata.\textsuperscript{27} One possible explanation of PIB is that individuals who have experiences and histories of hostile and otherwise negative interactions with their social worlds have developed readily accessible knowledge structures (or schemata) that represent members and events in their social environments to be intentionally dangerous and harmful. These knowledge structures affect social perception at several levels and may become highly organized and automatically activated as they are repeatedly accessed and reinforced over time. In effect, hostile-world schemata may set the SIP course, including interpretation and response accessing, for aggressive individuals.\textsuperscript{28} In this way, an aggressive individual’s SIP may become restricted in that, over time, he or she may become less able to interpret social situations as nonprovocative because hostile and negative world schemata are more readily—if not automatically—accessed and active in guiding social–cognitive and behavioral reactivity.

Also relevant is recent empirical research on college-aged adults regarding the relation between trait anger and attentional bias for angry faces. In their 2005 multieperiment investigation, Jack van Honk and his colleagues found that trait anger, as opposed to trait anxiety, was positively associated with being biased toward attending to angry faces.\textsuperscript{29} This attentional bias was further discerned to be present even on a preconscious level, suggesting that the connection between trait anger and biases that favor attention to angry faces is processed automatically, as opposed to conceptually. This set of findings may be relevant to

\textsuperscript{26} Id.
\textsuperscript{29} Jack van Honk, Adriaan Tuiten, Edward de Haan, Marcel van den Hout & Henderickus Stam, \textit{Attentional Biases for Angry Faces: Relationships to Trait Anger and Anxiety}, 15 COGNITION AND EMOTION 279, 279–297 (2001).
understanding the role of impulsivity in PIB in that being biased toward attending to angry faces may contribute to one’s tendency to quickly (and perhaps impulsively and automatically) interpret provocations in ambiguous situations. Given the well-established link between PIB operations and aggressive conduct problems, it is clear that this area of research needs to be extended to populations that vary in violent behavior.

Studies of the most popularly researched component of PIB, hostile attributional bias, suggest that there is a continuum upon which this type of cognitive functioning must be considered. Whereas some individuals may be only slightly more likely to attribute hostility to an ambiguous provocateur, others may be so biased in their interpretational processing that they attribute hostility in situations in which the evidence suggests that the actor’s intent is benign. Individual differences in hostile attributional processing may partly account for why alternative characterizations of this cognitive phenomenon have been preferred by different research scientists. For instance, since the term hostile attributional bias was first introduced and has become commonly used, other researchers have argued that the term is too strong and that hostile attributional style is more accurate; at the other end of the continuum, hostile attributional processing has been conceptualized as an actual cognitive deficit. A cognitive bias in its extreme form may stem from, or act as, an actual deficit in processing in that the person’s operating may be so distorted that he or she is literally unable to otherwise process information. In this way, extreme PIB has the psychologically restrictive effect of literally preventing one from adequately processing information and making valid attributions as to the meaning of social stimuli.

In summary, multiple cognitive processing difficulties may contribute to being biased toward interpreting social situations as provocative. Encoding deficits involve the failure to attend to relevant environmental information and social cues. In addition, individuals with encoding limitations may have difficulty organizing incoming information in ways that allow for accurate interpretation of social events. Attributional biases may involve the tendency to view social stimuli as intentional, harmful, hostile, and wrongful in their actions. Further problems with interpretation may include ascribing personally relevant meaning to social situations or strictly associating negative outcomes with the behaviors of the stimulus actor. Taken together, these processing difficulties are likely to contribute, in both additive and interactive incre-

---


ments, to the likelihood that a person would experience a situation that is ambiguous as to its provocative content as decidedly provocative. The opposite may be true, of course, with respect to someone who has cognitive biases that are otherwise directed. The individual who is biased in favor of benign or positive interpretations of his or her social world may be less likely to react with hostility, even when faced with an unambiguous provocation. As such, it is the encoding and interpreting of a social stimulus that allows one to understand and make meaning of it so that he or she can respond in kind. If a person is interpretationally biased in a particular direction, it is likely that this bias is reflected by his or her pattern of behavioral reactivity.

This relation between PIB and antisocial reactivity may be neatly applied to the reactive killer. PIB may preclude a person, in certain situations, from believing anything other than that he or she was egregiously provoked. In turn, this bias promotes the likelihood that the type of elevated fury that is associated with heat of passion violence would be experienced. In this way, one’s ability—or lack thereof—to validly interpret social cues is critical to discerning his or her mens rea in any case of reactive homicide (or any case of reactive violence, for that matter).36

Cognition’s Role in Attributing Wrongfulness in Reactive Homicide Cases

Homicidal crimes are defined and discerned, in significant part, by their respective mentes reae. A defendant charged with murder may invoke the heat-of-passion defense and demonstrate that, as a direct result of a substantial provocation, he or she became emotionally overwhelmed and killed the victim while in this uncontrolled state. This affirmative partial excuse defense reduces murder to manslaughter on the basis that the emotional charge negates the more serious mens rea that is required for murder. However, the criminal culpability of the reactive killer who has a cognitive deficiency in the form of PIB and, as a result, misinterprets an ambiguous situation to be substantially provocative, may not be similarly mitigated. In other words, despite the fact that the defendant who has dysfunctional processing may also lack the necessary mens rea for murder, his or her crime will not be reduced to manslaughter if the provocation in question does not satisfy the reasonable-person requirement. This is true even in the case of the defendant whose PIB is so severe that he or she is cognitively deficient and could not have otherwise interpreted the situation at hand.

This is morally problematic in that the rationality of the defendant in this latter instance is diminished in a manner that is outside of his or her control and is thus unpreventable. Surely, it is not the case that this killer is blameless; in fact, his or her crime is intentional and egregious. However, in line with doctrines of

36 Note that the heat of passion defense applies only to murder. Committing a nonhomicidal violent act in response to a substantial provocation and in the heat of extreme fury is not defendable by the heat of passion doctrine. However, the limited application of the heat of passion doctrine to reactive homicide is a separate problem (or set of problems), on both logical and moral grounds, as it is unclear why diminished rationality would translate into diminished responsibility for reactive killings but not other kinds of reactive violence. This issue demands a broader discussion that is not only beyond the scope of the present paper, but has been rigorously addressed elsewhere. For example, see Stephen J. Morse, Diminished Rationality, Diminished Responsibility, 1 Ohio St. J. of Crim. L. 289, 289–308 (2003).
retributive justice and penal proportionality, it is equally certain that this killer lacks the requisite mens rea of murder. As discussed, particularly in its extreme form, PIB serves to restrict a person’s ability to interpret an ambiguous provocation as anything but a provocation. As Stephen J. Morse has argued, “Diminished rationality does not simply function as a ‘but for’ cause of the criminal conduct. It must also have substantially impaired the defendant’s ability to access and consider reflectively the good reasons not to commit the crime on this occasion.” In fact, this is precisely how PIB may function in reactive violence; it may prevent the responding individual from accessing a good reason to not retaliate in that it may preclude an interpretation that the stimulus is anything other than seriously provocative.

Reactive killers whose interpretational processing is distorted vary, of course, in their moral (although not necessarily criminal) culpability based on, among other factors, the severity of processing dysfunction. They may also vary in their retributive judgment style. Neil Vidmar presented a model of retributive judgment by which (a) an individual perceives a rule or norm violation; (b) the individual judges the rule violator’s intention to be blameworthy; (c) the combination of Steps 1 and 2 offends the responding individual’s self, status, or values; (d) anger is elicited; (e) the cognitions and emotion described in Steps 1 through 4 cause behavioral retaliation; and (f) the respondent’s mental state returns to normal during or after the punishment and as it becomes evident to the respondent that retributive justice has been exacted. Although this model has not received the rigorous empirical testing that Crick and Dodge’s SIP model has, several elements of Vidmar’s model are strikingly similar to those of Crick and Dodge. In essence, both models articulate a cognitive response sequence by which an individual perceives a wrongful act, ascribes personal meaning to the act, is personally offended by the act, and responds with anger and retaliatory behavior in an attempt to “get back at” the perceived provocateur (or rule violator, in Vidmar’s terms).

Although the terminology may be different, the relevance of PIB to the respective models of Crick and Dodge and Vidmar is identical. In Vidmar’s model, the responding person may be biased in favor of perceiving a rule or norm violation in situations that are ambiguous in this respect. A person may be further prone to

39 Id.
41 Id.
43 Id.
attribute malicious or hostile intent to the rule violator in cases in which the actor’s intent is unclear. The responding individual may also be biased toward attributing personal meaning (e.g., that the rule violator’s actions are personally directed) to the witnessed event. Whereas these cognitive distortion patterns map onto the first three steps in Vidmar’s model, they are not substantively different from aspects of PIB that have already been discussed with respect to Crick and Dodge’s model of SIP. These cognitive processes may be activated and transpire in a matter of seconds, or even microseconds. Rapid (or impulsive) processing of these perceptual and evaluative judgment domains may be particularly characteristic of individuals who are severely biased in their cognitive operating.44 Dodge and I have argued that this type of biased, impulsive judgment processing serves to preempt fuller processing from occurring and may even preempt certain critical cognitive operations from being accessed and activated.45

Each of these scholarly traditions and literatures contributes to a more comprehensive understanding of social–cognitive foundations of antisocial behavior, and, more specifically, the relevance of biased cognitive processing. Translated into terms of criminal culpability, individuals with processing tendencies that are strictly biased and occur rapidly in real time are less able (or perhaps even unable) to process incoming social information in ways that promote nonaggressive, adaptive emotional and behavioral functioning. Having now covered a broader range of psychological processes in relation to reactive violence than has been addressed in related writings, we may turn to some additional issues that have implications for various areas of public policy.

The Relative Culpability of the Cognitively-Biased Heat-of-Passion Killer

As discussed, by current common law framings of heat of passion in United States common law,46 reactive killers who are dysfunctional in their interpretational processing may fail to successfully invoke the heat-of-passion defense if their interpretation of provocation falls short of the reasonable-person standard. The explanation for this distinction may lie in the fact that the heat-of-passion killer with PIB does not neatly fall into either category of murder or manslaughter, in at least three ways.

1. Absence of Both Substantial Provocation and Malice Aforethought

Consider the instance of the heat-of-passion killer with PIB who misinterprets a situation as substantially provocative and responds with fury and homicidal retaliation. Whereas this killer may lack the malice aforethought and premedita-

45 Id.
46 How Model Penal Code jurisdictions may judge a heat of passion case in which the defendant has demonstrated cognitive dysfunction consistent with PIB is another question. The “extreme mental or emotional disturbance” language of the Code is likely open to a much broader spectrum of application. See MODEL PENAL CODE § 210.3(1)(b) (1962). The current discussion has focused on the American common law construction of the heat of passion doctrine.
tion that is associated with murder, he or she has not actually been provoked, or perhaps, at least, not substantially so. The absence of reasonably interpreted, substantial provocation precludes the homicidal act from being partially excused. In fact, some legal scholars argue that the provocative act should not only be reasonably interpreted but also be both real and unlawful or even feloniously violent.47 One obvious benefit that would result from revising the doctrine in this manner is that provocation may be less subjectively defined—that is, sufficient evidence would be required that the provocative act in question was a violent felony, as defined by law. However, this definition rules out numerous examples of provocation that have historically been judged to be sufficient provocation per the reasonable-person requirement (including adultery, arson, etc.), and it has thus been conceded that restricting adequate provocation to violent felonies may be too restrictive.48 Of course, restricting the definition of adequate provocation to felonious or seriously violent acts eliminates the dilemma of PIB, because it makes the experience of the reactive killer moot in the case that he or she has been presented with no criminal provocation.

Whereas redefining adequate provocation in this manner may reduce the degree of subjectivity,49 it does not address the issue of mens rea. The person whose reactive homicide is based on his or her dysfunctional cognitive and emotional processing does not kill with the same premeditation and malice aforethought as does the murderer, even in cases in which the reactive killer kills intentionally. Furthermore, to the degree that the heat-of-passion doctrine is based on the principle of penal proportionality, it fails to resolve the issue that reactive killers with PIB do not perfectly fit into either category of manslaughter or murder. Questionable (if not outwardly unjust) though it may be, the reactive killer with PIB is classified as a murderer because, although he or she may well have provided evidence that negates the requisite mens rea of murder, the elements of the heat-of-passion defense are not satisfied in that adequate provocation is not demonstrated according to the reasonable-person requirement.

2. Future Dangerousness

One rationale for distinguishing heat-of-passion manslaughter from murder has less to do with retributive justice than it does with predicting the future. One who murders commits a premeditated killing in cold blood, typically to realize an instrumental goal. In contrast, the heat-of-passion killing, although intentional, is

49 Note that although restricting adequate provocation to acts that are illegal or felonious, or even to violent felonies, may reduce subjectivity in the process of determining whether the provocation at hand suffices, it does not eliminate subjectivity as the court would still need to discern the legal status of provocation in question.
a less mindful act in that the actor’s rationality and control are undermined by
intrusive emotional disturbance. This difference in degree of intentionality (or
mindfulness, or mens rea) has been argued to equate to a difference in future
dangerousness, or the likelihood that the defendant will kill again. However, the
idea that individuals who commit instrumental homicide (murder) differ in future
dangerousness from those who commit reactive homicide (heat-of-passion man-
slaughter) has not been empirically established; likewise, differential stability of
instrumental and reactive violent subtypes across time in general (i.e., beyond
homicidal violence) has not been scientifically substantiated.

In the possible case that this distinction is valid, however, the heat-of-passion
killer with PIB presents a unique case. This individual does not have the degree
of intentionality that is associated with murder and a greater likelihood of
homicidal recidivism but may be argued to be more likely to kill again than the
heat-of-passion killer who does not have PIB. This is because the person who is
biased toward interpreting ambiguous situations as provocative is likely to have a
larger set of social stimuli that may trigger a homicidal response. Here again, the
heat-of-passion killer with PIB does not fit neatly into either the murder category
or the manslaughter category.

Of course, the issue of future dangerousness is relevant in discerning levels of
homicide only to the degree that the guidelines of this endeavor stray from their
retributive foundations. The retributivist would insist that the degree of guilt
and severity of punishment that are ascribed to the killer be determined exactly
by the culpability of the act in question. Questions about the likelihood of acts
not yet committed (if they are to be committed at all) should not factor into
determinations of retributive justice. Still, future dangerousness continues to
be a consideration in judicial decision making, with respect to issues of crime
and punishment, that is not strictly bound by retributivist principles.

If the heat-of-passion defense were to be reframed to account for PIB, what
effect, if any, may this have on the debate about homicidal subtypes and future
dangerousness? As discussed, there is no scientific substantiation for the claim
that one violent subtype (instrumental vs. reactive, murder vs. manslaughter) is
more stable, and thus more predictable, than the other. As noted, there has been
significant debate as to the relative future dangerousness of the cold-blooded,
premeditated murderer as compared with the reactive, heat-of-passion killer, yet
such debate has not been empirically resolved. It is likely that insufficient

50 See Joshua Dressler, Why Keep the Provocation Defense?: Some Reflections on a Difficult
Subject, 86 MINN. L. REV. 959, 959–1002 (2002) (providing an analysis of the utilitarian rationale
for future dangerousness with respect to the heat of passion defense). Cf. JAMES STEPHEN, A HISTORY
OF THE CRIMINAL LAW IN ENGLAND (Vol. 3, 1883).

51 See Frank Vitaro, Mara Brendgen & Edward D. Barker, Subtypes of Aggressive Behaviors:

52 Although, for the sake of argument, a case could be made that merely presenting danger of
wrongdoing is itself wrongdoing and thus punishable.

53 See Stephen J. Morse, Diminished Rationality, Diminished Responsibility, 1 OHIO ST. J. OF
scientific investigation of this issue has been conducted; however, it may also be that the future dangerousness of instrumental and reactive forms of violence does not empirically differ (research science that produces statistically nonsignificant or null results is typically not published or otherwise publicly accessible). In addition to the issue of future dangerousness being out of place in a retributive rationale for provocation law, the fact that violent (or, more specifically, homicidal) subtypes have not been empirically demonstrated to have differential stability provides further justification to remove this consideration from the PIB heat-of-passion analysis, at least for the time being.

3. Provocation Interpretational Bias and Mental Disorder

In most states, murder may be reduced to manslaughter either because of satisfactorily meeting the elements of the heat-of-passion defense or by demonstrating that other mitigating circumstances compromised the defendant’s rationality at the time of the killing. Often, criminal culpability may be mitigated by presenting sufficient evidence that the defendant suffers from (or at least suffered from at the time of the act in question) an officially recognized mental (or psychiatric or medical) disorder (or diagnosis, disease, or defect), such as major depressive disorder or posttraumatic stress disorder (PTSD), or by showing that the defendant takes prescribed medication that can have side effects that complicate or hamper mental functioning—that is, the defendant may show that a mental disorder or prescribed medicine diminished his or her rationality at the time of the killing.

PIB is not an official mental disease or disorder; nor is the most heavily researched and discussed aspect of PIB, hostile attributional bias, an official mental disease or disorder. Still, there exists no legitimate rationale for which a real, scientifically substantiated, nonculpable mental disturbance that diminishes rationality but is not an official mental disease/disorder should not be accepted as a basis for mitigating criminal culpability.

Although PIB is not an official psychiatric diagnosis, it has been associated with several official mental disorders in the scientific literature. For example, individuals with borderline personality disorder (BPD) are typically more sensitive to minor stressors than are others. When presented with relatively minor provocations, they tend to become aroused more quickly and more dramatically. In addition, it normally takes individuals with BPD longer to cool down after becoming aroused in response to experiencing provocation. It has also been suggested that individuals with BPD who are subjected to child abuse may become particularly sensitive to cues of threat and become biased toward attributing hostility and intentional harm to others in situations that are unclear as to their social meaning. This developing impairment in SIP may contribute to an increasing likelihood that the individual with BPD retaliates more frequently,

---

54 The term “mental disease or defect” is used by the Model Penal Code. See Model Penal Code § 4.02 (1985). To be clear, the Model Penal Code’s “extreme mental or emotional disturbance” provision [Model Penal Code § 210.3(1)(b) (1962)], that consumes traditional heat of passion doctrine, is separate and does not make mention of mental disease or defect.

quickly, and severely to modestly or ambiguously provocative (or even unambiguously nonprovocative) stimuli, because of their distorted interpretations that such cues are significantly provocative or threatening.56

Some scholars have conceptualized BPD as a type of PTSD.57 PTSD may be diagnosed when an individual suffers a significant trauma by which he or she experiences substantial emotional disturbance (intense fear, helplessness, or horror) and develops a pattern of reliving the traumatic event, avoiding stimuli associated with the trauma, and experiencing increasing arousal.58 SIP problems have been found to partially mediate early physical maltreatment and later antisocial behavior in correlational research.59 Other studies have shown that posttraumatic stress symptoms and PTSD are positively related to attributional and interpretational biases and errors,60 suggesting that the potentially significant and pervasive effects of a trauma may play a role in distorting or even restricting the victim’s future ability to interpret social events that are open to being interpreted as provocative, threatening, or otherwise harmful.

Psychiatric disorders that are based, in part, in paranoia (e.g., paranoid schizophrenia, paranoid personality disorder) have also been associated with PIB. Paranoid individuals tend to believe that other people are intending or even scheming to cause them harm, even when there is little or no evidence in favor of such a belief. Paranoid personality disorder, for instance, is characterized, in part, by tendencies to suspect others of being exploitative, harmful, and deceitful and to distort experiences by interpreting nonprovocative cues as hostile and read “hidden demeaning or threatening meanings into benign remarks or events.”61 PIB, in its most extreme form, may present in a fashion that is similar to paranoia in that a person may be restricted to finding threatening meaning across varied social events that bear little to no resemblance to objective threats.62

PIB has also been linked to depression.63 Harold Leitenberg and his colleagues


57 See, e.g., Judith L. Herman, Trauma and Recovery (Basic Books 1992).


62 In addition, extreme PIB (or paranoia) may be relevant to imperfect self-defense, a separate partial-excuse defense that is discussed below.

(1986) found that depressive symptomatology was positively correlated with cognitive distortions of overgeneralizing the meaning of negative events and interpreting such events as catastrophic. Subsequently, in their 1992 study, Nancy L. Quiggle and her coinvestigators found that not only were aggressive children more likely to be biased toward attributing hostility to ambiguous social cues, but so were depressed children. Similar findings have been observed in incarcerated adults. In addition, Jennifer E. Vitale and her colleagues (2005) found that hostile attributional bias was positively related to a depressogenic attributional style in adult male offenders, suggesting that at least this one aspect of PIB may contribute to certain individuals’ tendency to negatively view the self, others, and the social world in general. It should be noted, however, that hostile attributional bias in some depressed individuals may be more likely to lead to withdrawal and self-isolating behaviors than to reactive violence. Still, these studies provide further evidence that PIB is related to psychiatrically disordered functioning that has been legally recognized (e.g., major depressive disorder) to diminish one’s rationality.

This discussion points to linkages between PIB and several official psychiatric disorders. Whereas these and other diagnoses may be recognized with respect to mitigation of criminal culpability, it remains unclear as to why the law’s focus is on official mental disorders as opposed to the specific mechanisms by which rationality is diminished. It seems more sensible to attend to the actual processes of cognitive dysfunction that potentially serve to impair rationality rather than to diagnostic classifications within which these processes may be included. The undiagnosed individual with PIB may be no less diminished in his or her rationality than the person who meets all symptom criteria (many of which may have no relevance to, or bearing on, one’s rational capacity) for an official mental diagnosis.

Additional Policy Implications of Research on Provocation Interpretational Bias

Legal Policy: The Case of Imperfect Self-Defense

It should also be recognized that there may exist other affirmative defenses to which PIB may be relevant. Perhaps most obvious to consider is imperfect self-defense. Self-defense may be raised as a defense to certain criminal charges and civil claims, with the most typical cases being ones of homicide, assault and battery, and attempts to exact violence against an individual. A claim of self-defense requires the

---

64 Harold Leitenberg, Leonard Yost & Marilyn Carroll-Wilson, *Negative Cognitive Errors in Children: Questionnaire Development, Normative Data, and Comparisons Between Children With and Without Self-Reported Symptoms of Depression*, 54 J. OF CONSULTING AND CLINICAL PSYCHOL. 528, 528–536 (1986). Note that this biased processing of overgeneralization and catastrophization of negative events may be compared with the process of “splitting” that is associated with borderline personality disorder by which a person views everything as either good or bad based on individual positive or negative experiences or events.


defendant to show that his or her actions were justified. Generally stated, a person may use force when it is reasonably necessary to prevent one’s own physical harm. The degree of force exerted by the defender may not be disproportionately more than the amount of force that he or she reasonably believes is presented by the threat (or attacker). Thus, whereas deadly force may be used in defense against a deadly attack, it may not be used in a situation in which the defendant is at risk of suffering nondeadly physical harm (excluding grievous bodily harm). Alternatively, imperfect self-defense may be claimed or found in situations in which self-defense was required but the defensive act was somehow unreasonable. In the case of the defendant who had a good-faith belief that the amount of defensive force used was necessary but his or her belief fails to reflect that of a reasonable person, the defendant’s criminal culpability may be mitigated but not fully exonerated. Although this defense is not as commonly recognized as heat of passion and some other affirmative defenses, a minority of jurisdictions do allow imperfect self-defense to reduce murder to manslaughter.

As with heat of passion, the success of an imperfect self-defense claim depends largely on the reasonable-person requirement. To the degree that the amount of defensive force used is determined to be unreasonable by the fact finder, the defendant’s criminal culpability is not mitigated. It is easy to imagine how PIB may be relevant to imperfect self-defense, in that the defendant with PIB may be restricted to interpret a situation as substantially harmful and, in accordance with this interpretation, enact an amount of reactive violence that is proportionate to his or her good-faith belief or interpretation of physical danger but is unreasonable in the eyes of the fact finder. As a result, the defendant with PIB may be found guilty of, and punished for, a crime for which he or she did not have the requisite mens rea.

Consider an example of someone who kills in response to having a realistic water pistol pointed at him or her. The fact finder may rightfully consider numerous factors in determining the reasonableness of the killer’s belief that his or her life was truly jeopardized: How realistic were the physical features of the water pistol? What was the relationship of the provocateur and the defendant? In what context was the water pistol drawn on the defendant? In the case of the defendant with PIB, the mere presence of a material object that resembles a gun may be enough to trigger a rigid interpretational tendency that produces a judgment that the “provocateur” intends to kill. Perhaps the defendant has a history of being threatened or attacked with a gun; perhaps he or she grew up in a hostile environment and was witness to a world of chronic violence; perhaps there is a less obvious course by which the defendant may have developed PIB. Regardless, the role of PIB in the defendant’s retaliation acts to justify, in the defendant’s mind, his or her violent response to the provocateur.

67 Justification defenses differ from excuse defenses in that the defendant’s actions are determined to have been warranted, given the circumstances. In excuse defenses, the defendant’s actions are agreed to have been wrongful, but the defendant’s criminal culpability is partially or fully excused because of individual-specific features (e.g., mental dysfunction) that are meaningful as to how and why the defendant acted as he did. As Austin clarified: “In the one defence, briefly, we accept responsibility but deny that it was bad; in the other, we admit that it was bad but don’t accept full, or even any, responsibility.” John L. Austin, A Plea for Excuses, in 57 Proceedings of the Aristotelian Society 1, 2 (Mark Eli Kalderon ed., 1956–1957).
Although the issue of PIB appears to raise a moral problem for imperfect self-defense, it does not suggest that this defense presents a logical inconsistency, as it does for heat of passion. Whereas heat of passion allows unreasonable behavior to be excused because of mental dysfunction that is emotional but not cognitive, imperfect self-defense makes no allowance for unreasonable behavior regardless of the nature of the defendant’s claimed mental dysfunction. This is because, unlike heat of passion, imperfect self-defense is not founded upon diminished rationality. For this reason, imperfect self-defense makes no allowance for PIB-related cognitive dysfunction, despite its clear relevance to the reactive nature of the violence to which imperfect self-defense applies. This distinction is the basis for why the present article focuses on heat of passion, as opposed to other affirmative defenses such as imperfect self-defense; PIB is directly relevant to diminished rationality and control, which lie at the very foundation of the heat-of-passion doctrine.

Mental Health and Criminal Justice Policy: Rehabilitation of Offenders

Numerous factors should be accounted for in the development of a comprehensive cognitive restructuring program for reactive violent offenders with PIB. By definition, individuals with serious PIB tend to judge ambiguous situations as provocative. They tend to make such evaluations quickly, and they are more inclined to judge the degree of provocation as more severe. Cognitive restructuring techniques with such individuals may only be useful to the degree that they can learn to slow down their interpretational processing when presented with a social stimulus. Rapid processing and behavioral reactivity to social cues may signify that a person is functionally bypassing stages of SIP (e.g., clarification of goals, generating different behavioral options, and evaluating and comparing alternative ways of responding) that are important to adaptive, nonviolent functioning. To the degree that processing speed may be reduced, individuals with PIB may be better able to consider alternative understandings of social events with which they are presented. Early steps of any cognitive restructuring program with individuals with PIB should focus first on processing speed and then on perspective taking.

Advanced rehabilitation for reactive violent offenders with PIB may be able to attend to finer aspects of processing. In addition to being biased toward interpreting ambiguous social situations as provocative, some individuals with PIB may be more likely to (a) judge real but modest provocations as substantial (i.e., exaggerate the severity of the provocation); (b) believe (or even insist) that an ambiguous provocateur’s intention is to do wrong and cause harm, even when he or she is aware of nonprovocative explanations for the given social event; (c) anticipate only harmful outcomes of the perceived provocation, which may contribute a disproportionately high degree of reprehensibility being attributed to the event; (d) personalize the provocation; and (e) take offense to provocations even in cases in which it is clear to the responding individual that the provocation is not directed at him or her. This last aspect of PIB may be observed in individuals who retaliate against provocateurs for the purpose of defending others whom they perceive to have been wrongly provoked or unfairly treated.

It has been suggested that the reactive aggressive individual has a less
promising prognosis than the instrumental aggressor, presumably because reactive aggression is founded in persistent biochemical, psychophysiological, and psychological functioning, whereas instrumental offending has been associated less with complex intrapersonal functioning and more with environmental contingencies. At present, this is an empirical question that remains unanswered. I suggest, however, that the opposite may be the case, largely because of the component of moral judgment that has been, at least conceptually, associated with reactive, but not instrumental, violence. Whereas the instrumental aggressor acts out of instrumental desire, the reactive aggressor retaliates because he or she has interpreted the actions of his or her victim to be provocative and, presumably, unjust. As such, it may be that the reactive offender with PIB, because of his or her moral cognizance, is more likely to regret his or her violent actions and be motivated to change so that he or she does not behave similarly in the future. Despite increasing scientific attention to instrumental versus reactive aggression during the past 20 years, relatively little research on moral–cognitive differences between aggressive subtypes has been conducted.

Conclusion

This article examines the murder–manslaughter distinction from a SIP perspective and discusses critical, previously neglected legal and policy implications for provocation law in the United States. In addition, this article illustrates the unavoidable imbalance of partially excusing criminal culpability due to emotional, but not cognitive, mental impairment. A uniquely broad discussion of the relevant psychological literatures is provided, and a three-class analysis of how the law may fail to equally apply to reactive homicide offenders is advanced. Finally, applications of cognitive bias to other legal (e.g., imperfect self-defense) and mental health (e.g., rehabilitation of offenders) policy issues are presented. This article provides further evidence in favor of the argument that, in the case of the reactive killer with significant interpretational bias, there is no legitimate reason why his or her criminal culpability should not be mitigated in the same way in which the heat-of-passion doctrine currently provides for emotionally dysfunctional killers who commit homicide in response to reasonably judged (or “adequate”) provocation.

Received June 16, 2008
Revision received August 5, 2008
Accepted August 11, 2008