

# “Blunt Not the Heart, Enrage It””

The Psychology of Revenge and Deterrence



Rose McDermott, PhD | Anthony C. Lopez, PhD | Peter K. Hatemi, PhD

Texas National Security Review: Volume 1, Issue 1 (December 2017)

Print: ISSN 2576-1021 Online: ISSN 2576-1153

Malcolm  
 Dispute it like a man.  
 Macduff  
 I shall do so;  
 But I must also feel it as a man:  
 I cannot but remember such things were,  
 That were most precious to me. — Did heaven look on,  
 And would not take their part? Sinful Macduff,  
 They were all struck for thee! naught that I am,  
 Not for their own demerits, but for mine,  
 Fell slaughter on their souls: heaven rest them now!

Malcolm  
 Be this the whetstone of your sword. Let grief  
 Convert to anger; blunt not the heart, enrage it.

- *Macbeth*, William Shakespeare

Why is the instinct for vengeance so strong even when it is clear that widespread death and destruction would be a much more likely outcome than any kind of “victory”? In the event of a nuclear war, why is second-strike retaliation so certain when it may gain nothing of social or material value? We believe these things because humans share a universal thirst for retaliation in the face of threat and in the wake of loss, no matter what classical economists may say to the contrary about how people “should” behave. Indeed, the psychology of revenge and the hatred on which it rests make a seemingly irrational second strike entirely credible. We can apply this analysis to nuclear weapons, but the basic drive is no different than the one that makes most people want to kill anyone who threatens their child, or to hurt a cheating spouse. The instinct for revenge is universal, automatic, and immediate. It also serves a function: to deter

the threat of future exploitation.

As long as humans have lived and competed in groups, the question of deterring threats from one’s adversaries has been of central importance. Humanity’s progression from living in small hunter-gatherer tribes where everyone knew one another to nation-states with millions of people has, in many cases, magnified the stakes of the challenge rather than altered its fundamental dynamics.

For all of human history, people have had to deal with challenges to their physical security and that of their family and friends. Aggression as an adaptation for conflict resolution has been extensively studied in primates<sup>1</sup> and in humans.<sup>2</sup> What has broadly been labeled “retaliatory aggression” (most often immediate but also delayed) is one of the most zoologically common, well-recognized, and well-studied behavioral responses for dealing with threats and challenges.<sup>3</sup>

1 James Silverberg and J. Patrick Gray, *Aggression and Peacefulness in Humans and Other Primates* (New York: Oxford University Press, 1992); Richard Wrangham and Dale Peterson, *Demonic Males: Apes and the Origins of Human Violence* (Boston: Houghton Mifflin Harcourt, 1996).

2 Martin Daly and Margo Wilson, *Homicide* (New Brunswick: Transaction Publishers, 1988); Aaron Sell, John Tooby, and Leda Cosmides, “Formidability and the Logic of Human Anger,” *Proceedings of the National Academy of Sciences* 106, no. 35 (September 2009): 15073–78.

3 John Archer, *The Behavioural Biology of Aggression* (Cambridge: Cambridge University Press, 1988); J. Martin Ramirez and Jose M. Andreu, “Aggression’s Typologies,” *International Review of Social Psychology* 16, no. 3 (2003): 125–41.

Why not then also recognize that “revenge” specifically is located within the evolutionary logic of retaliatory aggression more broadly? Indeed, we find that it evolved because of its ability to solve the recurrent challenge of deterrence, which has existed throughout the human experience and has clear implications for reproductive fitness.<sup>4</sup>

How can we know this? Scholars of human behavior often begin with our closest evolutionary cousins — chimpanzees and bonobos — and look for contrasts and parallels between these species and our own. Primate research has revealed that retaliatory aggression is a trait we undeniably share with non-human primates. For example, both chimpanzees and bonobos exhibit propensities toward individual and group-level retaliation, suggesting that the tendency toward retaliatory aggression dates at least to our most recent common ancestor approximately 5 to 7 million years ago.<sup>5</sup> In other words, retaliatory aggression in humans can be at least partly explained as a component of an evolved psychology we share with our primate ancestors.<sup>6</sup>

Conventionally, we say that deterrence is successful when the threat of unacceptable costs prevents an adversary from taking some undesired course of action. When effective, deterrence can achieve policy goals on the cheap and can mitigate the potential for unwelcome blowback. Failures of deterrence, however, can lead policymakers to throw good money after bad and to engage in reckless brinkmanship.

One of the best examples comes from Richard Nixon, who used such logic to seek an end to U.S. military involvement in Vietnam. One day, walking along a fog-shrouded beach in California, he told Bob Haldeman, his chief of staff:

I call it the Madman theory, Bob. I want the North Vietnamese to believe I've reached the point where I might do anything to stop

the war. We'll just slip the word to them that, “for God's sake, you know Nixon is obsessed about Communism. We can't restrain him when he's angry — and he has his hand on the nuclear button” — and Ho Chi Minh himself will be in Paris in two days begging for peace.<sup>7</sup>

While this strategy did not appear to work for Nixon, he believed that it would. As with Thomas Schelling's threat that leaves something to chance, or his notion of the rationality of irrationality,<sup>8</sup> Nixon believed that creating a reputation for disproportionate response would advantage his play against an adversary by encouraging it to back down in the face of threat.

Although scholars have developed an understanding of the strategic function of deterrence, we have a poor understanding of the psychological underpinnings of deterrence as well as the conditions under which deterrence is likely to succeed or fail.

Classic theories of deterrence emerged in the wake of the nuclear revolution and required that for deterrence to be stable, both actors had to commit to an otherwise seemingly irrational course of action: nuclear retaliation in response to a first strike.<sup>9</sup> Such a commitment is awkward within a rationalist framework because, as many theorists have pointed out, a second-strike attack cannot undo or mitigate the apocalyptic damage delivered in a first strike.<sup>10</sup> Despite the reluctance with which rational actors should commit to nuclear retaliation, history is replete with policymakers who have credibly and sometimes eagerly committed to just this course of action. Where *homo economicus* demands ambivalence at best, *homo sapiens* prove eager and ready.

We argue that the human psychology of revenge explains why and when policymakers readily commit to otherwise apparently “irrational” retaliation.

4 Michael E. McCullough, Robert Kurzban, and Benjamin A. Tabak, “Cognitive Systems for Revenge and Forgiveness,” *The Behavioral and Brain Sciences* 36, no. 1 (2013): 1-15.

5 Christopher Boehm, “Retaliatory Violence in Human Prehistory,” *British Journal of Criminology* 51, no. 3 (May 2011): 518-34.

6 Joseph H. Manson and Richard W. Wrangham, “Intergroup Aggression in Chimpanzees and Humans,” *Current Anthropology* 32, no. 4 (1991): 369-90; Azar Gat, *War in Human Civilization* (Oxford: Oxford University Press, 2006); Richard W. Wrangham and Luke Glowacki, “Intergroup Aggression in Chimpanzees and War in Nomadic Hunter-Gatherers: Evaluating the Chimpanzee Model” *Human Nature* 23, no. 1 (2012): 5-29.

7 H.R. Haldeman with Joseph DiMona, *The Ends of Power* (New York: Times Books, 1978), 83.

8 Thomas C. Schelling, *The Strategy of Conflict* (Cambridge: Harvard University Press, 1960).

9 Alexander L. George and Richard Smoke, *Deterrence in American Foreign Policy: Theory and Practice* (New York: Columbia University Press, 1974); Glenn Herald Snyder and Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises* (Princeton: Princeton University Press, 1977); Schelling, *The Strategy of Conflict*; Thomas C. Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966); Frank C. Zagare, “Reconciling Rationality With Deterrence: A Re-Examination of the Logical Foundations of Deterrence Theory,” *Journal of Theoretical Politics* 16, no. 2 (April 2004); Frank C. Zagare, “Rationality and Deterrence,” *World Politics* 42, no. 2 (1990).

10 Christopher H. Achen and Duncan Snidal, “Rational Deterrence Theory and Comparative Case Studies,” *World Politics* 41, no. 2 (January 1989): 143-69; Robert Jervis, “Deterrence Theory Revisited,” *World Politics* 31, no. 2 (January 1979): 289-324; Robert Jervis, “Rational Deterrence: Theory and Evidence,” *World Politics* 41, no. 2 (1989): 183-207; Robert Powell, *Nuclear Deterrence Theory: The Search for Credibility* (Cambridge: Cambridge University Press, 1990).



Indeed, we suggest that revenge offers the quintessentially, *psychologically* rational response to aggression. Revenge has several psychological attributes that are relevant for understanding deterrence. For example — and perhaps counterintuitively — revenge is not motivated by the rational expectation of future deterrence. It is instead driven by the intrinsic pleasure that one expects to experience upon striking back. The psychophysiological basis of this pleasure has been well-studied, and we understand that this internal reward system is designed precisely to distort cost-benefit analysis in adaptively useful ways.<sup>11</sup> It is precisely when revenge is sought for its own sake that it can be such an effective deterrent to adversaries and why it remains such an effective psychological strategy.<sup>12</sup> Revenge has evolved in part because of its deterrent effects, and these effects are greatest when retaliation is sought to satisfy a thirst for it, rather than as a product of conscious, time-consuming deliberation. For example, when someone catches a spouse cheating, particularly with a good

friend, he or she may desire revenge no matter the consequences. And most people would consider someone who stepped back and, before acting, rationally considered the costs associated with losing the marriage and friendship a bit odd or weak, even if that might be the more objectively rational strategy. An evolutionary perspective reminds us that it is important not to confuse the conscious or “proximate” goals of the actors (revenge) with the evolutionary or “ultimate” function of the evolved psychology behind it (deterrence).

The psychology of revenge is irrevocably embedded in notions of deterrence. Without such a foundation, no one would find the threat of retaliatory strike credible. But with the universal recognition of the automatic satisfaction that comes with revenge, few doubt that vengeance could very well lead to mutual annihilation. This helps to explain why policymakers are often willing to commit to a course of action that otherwise appears objectively irrational. Beyond identifying an evolutionary explanation for commitments

11 Robert H. Frank, *Passions Within Reason: The Strategic Role of the Emotions* (New York: W.W. Norton & Co., 1988); Dominique J. F. De Quervain, Urs Fischbacher, et al., “The Neural Basis of Altruistic Punishment,” *Science* 305, no. 5688 (August 2004): 1254.

12 Michael E. McCullough, Robert Kurzban, and Benjamin A. Tabak, “Putting Revenge and Forgiveness in an Evolutionary Context,” *Behavioral and Brain Sciences* 36, no. 1 (2013): 41-58.

to costly retaliation, we also offer a theoretically rigorous examination of revenge that is careful to distinguish it from other forms of retaliation. Negative reciprocity, for example, follows more of the tit-for-tat or an eye-for-an-eye kind of logic.<sup>13</sup> It can be cold and calculating, and seemingly more objective and proportional. Revenge is more of a psychological and emotional state that gets activated automatically and provides a strong drive in people who feel they have been wronged by another. It serves a deterrent purpose for the reasons laid out in greater detail below. People are not always driven by revenge when they retaliate. Still, revenge can feel really good when it is successful. By recognizing that different motivations can precipitate various retaliatory styles, we help to clarify the conditions under which policies of deterrence can lead to stable containment or destabilizing brinkmanship.

Leaders need not, and often do not, recognize the motivational distinction between those seeking revenge and those retaliating out of rational anger. Even when they are aware of a distinction, they may conclude that their adversaries are revenge-driven and hateful when some may not be, possibly losing important opportunities for avoiding conflict and achieving compromise. A fuller theoretical exposition of the meaning and function of revenge is central to understanding when and how conflict can be deterred. Deterrence, whether nuclear or other kinds, rests on the implicit assumption that the motive for retaliation is strong enough that, even when no benefit can accrue from launching a counterattack, the opponent should count on it anyway, and this belief will deter the initial assault.<sup>14</sup> Clearly a second-strike attack is not economically rational because it cannot prevent apocalyptic damage already sustained. Yet the universal recognition and appeal of the desire, and emotional pleasure, of revenge is part of what makes the threat of a second strike in a nuclear exchange credible.

We develop this argument as follows: First, we discuss the current foundations of deterrence theory in international relations. We then explain how psychological adaptations — which evolved

in the context of small-scale, hunter-gatherer communities rarely larger than 150 individuals — manifest within the modern environment of mass politics, particularly in the realm of nuclear deterrence. Our third section outlines the psychology of revenge from an evolutionary perspective and discusses how this might emerge in the context of modern conflicts. We explain that revenge evolved in part to respond to challenges and threats that required deterrence. Furthermore, we distinguish revenge from other forms of retaliation, such as negative reciprocity. Fourth, we discuss how particular emotions such as anger or hate can motivate revenge and other retaliatory possibilities.

## Why should individuals be so spiteful in the face of a threat that renders victory or redemption implausible, or death certain?

In the fifth section, we discuss how different contexts can trigger or mitigate various forms of retaliation. We then consider the implications for individual versus group-level analysis.<sup>15</sup>

### Nuclear Deterrence, Terrorism, and Revenge

The problem of deterrence is not unique to the modern international system, nor is it confined to the realm of nuclear strategy.<sup>16</sup> The emergence of nuclear weapons certainly precipitated a large wave of scholarship devoted to understanding the

13 Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

14 Schelling, *The Strategy of Conflict*; Schelling, *Arms and Influence*; Glenn Herald Snyder, *Deterrence and Defense: Toward a Theory of National Security* (Princeton: Princeton University Press, 1961); Zagare, "Reconciling Rationality With Deterrence: A Re-Examination of the Logical Foundations of Deterrence Theory"; Zagare, "Rationality and Deterrence."

15 In all of this, we do not claim that states are analogous to individuals. Rather, we argue that states are made up of individuals who share an evolved psychology of revenge that emerges in predictable ways within the context of the institutions that those individuals design.

16 Schelling, *The Strategy of Conflict*; Robert Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon* (Cornell: Cornell University Press, 1989); George H. Quester, *Deterrence Before Hiroshima: The Airpower Background of Modern Strategy* (New Brunswick: Transaction Books, 1986).

nature of deterrence and its transformation in the nuclear age.<sup>17</sup> Importantly, however, the underlying concept applies equally well to nuclear deterrence, conventional deterrence between states, and threats from individuals and other non-state actors, as well as between and among individual actors. Historically, successful deterrence rested on a state's ability to convince adversaries that it could deny their aims via conventional force of arms. In the nuclear age, however, deterrence is no longer a function of the conventional ability of armies to defeat armies. Instead, it is a function of a state's ability to deliver a similarly severe punishment to its opponent, even if the opponent is much stronger conventionally and *even after its own assured defeat*. Scholars classically identify a key attribute of the nuclear revolution as the shift from "deterrence by denial" to "deterrence by punishment." In this world, deterrence holds when threats to use nuclear weapons are credible, when neither side can hope to eliminate the other's retaliatory capabilities, and when the retaliation that is likely to follow any attack imposes a cost that is unacceptably high for each side.<sup>18</sup>

Regardless of the many ways nuclear weapons constitute a qualitative difference in weapons development, the crucial element that sustains effective deterrence is not so much the speed of nuclear destruction but, rather, assurance of the capability to deliver "mutual kill." In this way, the defender can destroy the attacker as likely as the reverse. This is what led President John F. Kennedy<sup>19</sup> to say after the Cuban missile crisis that "[w]e will not prematurely or unnecessarily risk the costs of worldwide nuclear war in which even the fruits of victory would be ashes in our mouths." It is what led President Ronald Reagan<sup>20</sup> to agree with Soviet premier Mikhail Gorbachev that "a nuclear war could never be won and must never be fought." Perhaps most presciently, Winston Churchill remarked that in the nuclear age "safety will be the sturdy child of terror, and survival the twin brother of annihilation."

Despite the horrors and intensity of these international dynamics, much of the literature on nuclear deterrence rests on a purely cognitive notion of credibility, which almost entirely excludes emotional foundations and motivations. This leaves out an important characteristic upon

which the edifice of deterrence depends. In short, despite arguments and assumptions that deterrence rests on assumed calculated rationality, the only truly credible aspect of deterrence lies in the authentic emotional power and psychological persuasion of the human drive for revenge in the face of violation or attack.

An evolutionary approach provides a set of tools for illuminating the emotional foundations of deterrence that are often assumed to be exogenous or are simply missing from the broader literature.<sup>21</sup> What benefit is there for the fallen in delivering a devastating post-mortem counterattack upon the assailant or to guarantee death by engaging in terrorist acts? This puzzle within the logic of nuclear and modern deterrence can be explained as a political manifestation of the human psychology of revenge. Specifically, as we discuss in detail below, the instinctual desire for revenge in response to a massive first strike is an important psychological foundation from which a credible threat to launch a retaliatory second strike can emerge, even after catastrophic defeat and death are assured.

Why should individuals be so spiteful in the face of a threat that renders victory or redemption implausible, or death certain? That is what our theory seeks to explain. This theory rests on a biological and psychological foundation of revenge, which feels so good that it overrides the cost-benefit analysis that would otherwise make people think before they act. And the near-universal recognition of the desire to give in to emotion at the expense of a more objective rational calculation under duress supports the credibility of a second-strike retaliatory deterrent threat.

First, the logic of modern deterrence rests fundamentally on the promise of revenge that can have the effect of altering adversary preferences *ex ante* by raising the prospect of unacceptable loss *ex post*. The notion of "retaliation" is endemic throughout the nuclear deterrence literature. The form that this promised retaliation takes is often assumed to be massive and disproportionate rather than gradual, which highlights the centrality of revenge. The classical deterrence literature has also emphasized that in the nuclear era, adversary intentions matter more than adversary capabilities. For example, Patrick Morgan presents the issue succinctly:

17 Robert Jervis, *The Illogic of American Nuclear Strategy* (Cornell: Cornell University Press, 1984); Schelling, *Arms and Influence*; Snyder, *Deterrence and Defense*; Bernard Brodie, *Strategy in the Missile Age* (Santa Monica: Rand Publishing, 1959).

18 Snyder, *Deterrence and Defense*.

19 John F. Kennedy, "Radio and Television Report to the American People on the Soviet Arms Buildup in Cuba," (October 22, 1962).

20 Ronald Reagan, "Address Before a Joint Session of Congress on the State of the Union," (January 25, 1984).

21 For an exception, see Jervis, *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon*.

Deterrence is undoubtedly a psychological phenomenon, for it involves convincing an opponent not to attack by threatening it with harm in retaliation. To “convince” is to penetrate and manipulate the thought processes of the opposing leaders so that they draw the “proper” conclusion about the utility of attacking.<sup>22</sup>

This representation of nuclear deterrence, with which we agree, may nevertheless be misleading if it guides some to the false conclusion that rational thought undergirds nuclear deterrence more than revenge. It is the emotional arousal resulting from the implacable willingness to inflict maximum physical damage on an adversary that, once demonstrated, inspires adversaries to halt. In addition, as noted above, revenge seeks suffering without understanding. Its goal lies in the elimination of the adversary because, correctly or not, prospects for future cooperation have been deemed impossible. In other words, a blind desire to cause suffering *regardless of what anyone thinks* has precisely the effect of changing what the audience thinks.

This dark desire reveals much about the logic of nuclear deterrence. In a world of nuclear-capable actors, a rational retaliatory form of harm should be unachievable because it could easily ignite a race toward first strike, since all would know that none should retaliate. This is far from strategic reality. Instead, nuclear deterrence is regularly cited as a defining element of the “great-power peace.” Indeed, deterrence can be stable, and it often is, because human actors automatically and universally recognize the plausibility of nuclear vengeance. The logic of revenge is further manifested in the nature of nuclear weaponry: They are primarily counter-value military instruments, designed to hurt people and destroy infrastructure, not primarily to target opponents’ nuclear weapons facilities.<sup>23</sup> They are taken seriously because of the speed and scale of the damage they can cause and because even weak opponents can harness this power to devastating effect.

This recognition raises another important point resulting from the desire to impose suffering without regard to cost. As Schelling<sup>24</sup> noted, the nuclear retaliation upon which deterrence depends is given special weight when it is imbued with an “automaticity” that is designed to remove deliberation and pause from the process of retaliation. This is famously illustrated in the movie “Dr. Strangelove,” in which a doomsday device makes retaliation automatic and irreversible. Of course, as the film brilliantly shows, a device guaranteeing destruction cannot serve a deterrent purpose unless the adversary is aware it exists, illustrating the psychological structure upon which the edifice of deterrence depends.

This example also highlights the problem introduced by the possibility that deterrence will fail if intentions and consequences are not fully and clearly communicated in advance. In addition, automaticity remains distinct from the invulnerability of a retaliatory response. An enemy must believe that a target’s force will survive an initial attack if a guarantee for retaliation is to remain credible. If there is any doubt that retaliation will not be automatic, dependable, and irreversible (and often disproportionate), nuclear deterrence becomes less than airtight.<sup>25</sup> Such assurance does not require policymakers to be hateful (although they often are) and it does not require policymakers to seek suffering for its own sake (although they often do). It requires only that the policy itself contain the recognizable attributes of vengeance: guaranteed, irreversible, disproportionate, and automatic retaliation. For example, any policy that promises irreversible and disproportionate violence in response to the crossing of a red line, especially when that line entails no real threat to one’s identity or welfare, may lack credibility in the eyes of adversaries. A recent example is President Barack Obama’s “red line” rhetoric regarding the use of chemical weapons in Syria, which produced no consequences once violated.<sup>26</sup>

This factor is what contributes to making the challenge of extended deterrence so vexing. It is hard to make the promise of retaliation credible in

22 Patrick M. Morgan, “Saving Face for the Sake of Deterrence,” in *Psychology and Deterrence*, ed. Robert Jervis et al. (Baltimore: Johns Hopkins University Press, 1985), 125.

23 Kenneth N. Waltz, “Nuclear Myths and Political Realities,” *American Political Science Review* 84, no. 3 (September 1990); Paul H. Nitze, “Deterring Our Deterrent,” *Foreign Policy*, no. 25 (Winter 1976-77). For challenges to this notion, see Austin Long and Brendan Rittenhouse Green, “Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy,” *Journal of Strategic Studies* 38, nos. 1-2 (2015).

24 Schelling, *Arms and Influence*.

25 It is also possible, of course, that a country would be deterred from a massive nuclear first strike simply by the threat of only a few of their cities being destroyed, i.e., the notion of “minimum” or “existential” deterrence. See Bradley S. Klein, *Strategic Studies and World Order: The Global Politics of Deterrence*, Vol. 34 (Cambridge: Cambridge University Press, 1994).

26 Greg Jaffe, “The Problem With Obama’s account of the Syrian red-line incident,” *The Washington Post*, October 4, 2016, [https://www.washingtonpost.com/news/post-politics/wp/2016/10/04/the-problem-with-obamas-account-of-the-syrian-red-line-incident/?utm\\_term=.0490d979236b](https://www.washingtonpost.com/news/post-politics/wp/2016/10/04/the-problem-with-obamas-account-of-the-syrian-red-line-incident/?utm_term=.0490d979236b).

the absence of sufficient emotional motivation for revenge. It also explains, at least in part, why some states put token groups of advisers as tripwires in allied territory:<sup>27</sup> Their presence does not make it more likely that a defensive operation would succeed. Rather, the prospect of their death ensures the emotional commitment designed to spark

## Inter-state peace stands perched upon the knife of vengeance.

the revenge-driven war upon which the credible threat of deterrence is based. This also heightens the strategic military importance of policies and procedures designed to enhance emotional and cultural connections between allied countries. After all, deterrence requires one of two elements to be successful: first, a truly vengeful policymaker (which cannot always be known a priori) or, second, a policymaker willing and able to “tie one’s hands” in a way that imbues state policy with the hallmarks of human vengeance: dependability, automaticity, irreversibility, and disproportionality.

Under conditions of maximum threat, a leader is most likely to be deterred when he believes that the cost of “guaranteed vengeance” from an enemy is too great to instigate an attack from the outset. While this is typically discussed in a nuclear context, it is no less true when attempting more conventional or personal deterrence. This requires that the opponents’ forces have a credible probability of surviving a first strike, as well as the belief that the adversary remains sufficiently vengeful to launch a counterattack even after absorbing a decisively destructive strike. Thus, although revenge is designed to weaken adversary capabilities, it may, as a byproduct, have the beneficial effect of recalibrating adversary preferences. As a result, deterrence theorists recognize that one of the best ways to alter adversary preferences is by paradoxically appearing to be blind to adversary preferences altogether and simply promising total destruction.<sup>28</sup>

Strategic nuclear weapons are obviously designed to inflict suffering on an adversary. As observed earlier, they empower losers to inflict retaliatory suffering *even after (state) death*. As Schelling<sup>29</sup>

notes: “Victory is no longer a prerequisite for hurting the enemy.” Nuclear weapons flatten the international hierarchy. Avoidance is not possible. Inter-state peace stands perched upon the knife of vengeance. In this way, advanced technology reduces or eliminates underlying power asymmetries that might otherwise be based on human strength, skill, intelligence, ingenuity, or other factors. Nuclear weapons represent the culmination of technological changes that have leveled pre-existing power asymmetries not only between states but also between individuals and states. Although the instinct for vengeance may be rooted in our evolutionary past, previously only states could amass the power to exert decisive state destruction. With nuclear weapons, individuals or small groups who can get ahold of such materials can also create massive damage. And both individuals and leaders acting on behalf of states share the basic psychological inclination for revenge in response to attack. Changes in weapons technology have made the role of revenge in providing the underlying emotional and psychological assurance of retaliation even more potent for establishing a credible deterrent.

However, the bond between revenge and nuclear deterrence is not immutable, and it can be broken. The greatest threats faced by nuclear states come in two forms; both represent a failure of deterrence. One is the possibility of a nuclear device in the hands of a political actor that has no “return address,” such as a terrorist group or criminal syndicate. Such groups are buffered from retaliation to the extent that they can simply disperse or move across state boundaries to avoid harm, or because they espouse an apocalyptic belief system whereby they do not care about, or fear, consequences that might result in their own death. They can effectively “run and hide,” a resuscitation of the avoidance mechanism once relatively available to our nomadic ancestors. In a world where deterrence depends on the certainty of retaliation, individual or non-state terrorist actors whose center of gravity is diffuse may be able to escape the constraints that would otherwise be imposed by their adversary’s expected retaliation. In these cases, retaliatory impulses will fuel rather than deter conflict.

The second class of threats is the existence of “rogue” actors impervious to the sort of incentives that would keep otherwise rational survival-minded state leaders in check. What makes

27 D. Bandow, *Tripwire: Korea and U.S. Foreign Policy in a Changed World* (Washington, DC: Cato Institute, 1996).

28 Thomas C. Schelling, *Micromotives and Macrobehavior* (New York: W.W. Norton & Co., 2006).

29 Schelling, *Arms and Influence*, 22.

rogue states and individuals dangerous, and what separates them from the type of coalitional threats encountered in ancestral environments, is not that they are vengeful or even suicidal but that they are more likely in the nuclear era to be able to deliver harm *despite* state suicide. A leader with an “enemy image” is often enough to undo the revenge-deterrence link, sometimes for good reason.

Both of these conditions threaten to break the bond between revenge and deterrence. This is not to say that leaders of rogue states are not “survival-minded.” A leader need not necessarily have suicidal intent to prove catastrophically destructive in the nuclear age. Some leaders, however rare, appear unwilling to give up power even when the alternative seems to ensure their own death. We now know, for example, that Fidel Castro intended to pursue a suicidal and preemptive nuclear attack against the United States had he been given authority over Soviet nuclear missiles during the Cuban missile crisis.<sup>30</sup>

## **Evolution and Security**

Given the pernicious and perennial centrality of revenge for deterrence, particularly in the nuclear age, what explains why humans possess this set of motivations? For a social species such as our own, challenges to survival and reproduction come not only in the form of harsh environments, scarce resources, and animal predation but, most importantly, in the form of conflicts with other individuals and competing groups. Evidence is accumulating that the long evolutionary history of humans living in groups has resulted in a complex “coalitional psychology,” which operates across almost all social dynamics, from cooperation and sharing to competition and aggression.<sup>31</sup> Indeed,

our ancestors were highly social, coalition-dwelling creatures for millions of years, and natural selection has shaped our minds accordingly.<sup>32</sup> What we call “retaliatory aggression” is a zoologically common, well-recognized, and well-studied behavior, designed by natural selection to deal with the challenges that threaten inherently social species.<sup>33</sup> Aggression as an adaptation for conflict resolution has been well studied in humans<sup>34</sup> and can be at least partly explained as a common component of an evolved psychology, refined over millions of years of human evolution.<sup>35</sup>

In order to further specify the nature and causes of retaliatory aggression in an international context, we introduce a framework that builds on and reconciles research in political science, psychology, and anthropology. All forms of retaliatory aggression can be placed on a behavioral continuum. We treat “retaliatory aggression” as the overarching or superordinate category, and “revenge” as an extreme form of retaliatory aggression. In other words, all revenge is an example of retaliatory aggression, but not all retaliatory aggression is properly considered revenge. As we consider it here, revenge is a specific form of retaliatory aggression that evolved at least partly because of its ability to solve the recurrent challenge posed by the adaptive problem of deterring adversaries.<sup>36</sup> A different and equally important form of retaliatory aggression is negative reciprocity, which we discuss in greater detail below. Distinguishing revenge from other forms of retaliatory aggression gives us greater insight into the nature of revenge and how it operates to produce deterrence. Furthermore, our argument about revenge supplies a necessary motivational dimension that is often missing from the literature on use of force.<sup>37</sup>

30 Sergo Mikoyan and Svetlana Savranskaya, *The Soviet Cuban Missile Crisis: Castro, Mikoyan, Kennedy, Khrushchev, and the Missiles of November* (Stanford: Stanford University Press, 2012).

31 John Tooby and Leda Cosmides, “Groups in Mind: The Coalitional Roots of War and Morality,” in *Human Morality and Sociality: Evolutionary and Comparative Perspectives*, ed. Henrik Høgh-Olesen (New York: Palgrave Macmillan, 2010); Anthony C. Lopez, Rose McDermott, and Michael Bang Petersen, “States in Mind: Evolution, Coalitional Psychology, and International Politics,” *International Security* 36, no. 2 (Fall 2011); Melissa M. McDonald, Carlos David Navarrete, and Mark Van Vugt, “Evolution and the Psychology of Intergroup Conflict: The Male Warrior Hypothesis,” *Philosophical Transactions of the Royal Society B* 367, no. 1589 (January 2012); M.B. Petersen and L. Aaroe, “Is the Political Animal Politically Ignorant? Applying Evolutionary Psychology to the Study of Political Attitudes,” *Evolutionary Psychology* 10, no. 5 (December 2012).

32 Frans De Waal and Alexander Harcourt, “Coalitions and Alliances: A History of Ethological Research,” in *Coalitions and Alliances in Humans and Other Animals*, ed. Alexander Harcourt and Frans B.M. de Waal (New York: Oxford University Press, 1992); Matt Ridley, *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* (New York: Viking, 1997).

33 John Archer, *The Behavioural Biology of Aggression* (Cambridge: Cambridge University Press, 1988); J. Martin Ramirez, Andreu Rodríguez, and José Manuel, “Aggression’s Typologies,” *International Review of Social Psychology* 16, no. 3 (2003).

34 Sell, Tooby, and Cosmides, “Formidability and the Logic of Human Anger”; Daly and Wilson, *Homicide*.

35 Gat, *War in Human Civilization*; Joseph H. Manson et al., “Intergroup Aggression in Chimpanzees and Humans [and Comments and Replies],” *Current Anthropology* 32, no. 4 (1991); Wrangham and Glowacki, “Intergroup Aggression in Chimpanzees and War in Nomadic Hunter-Gatherers.”

36 McCullough, Kurzban, and Tabak, “Cognitive Systems for Revenge and Forgiveness.”

37 Schelling, *Arms and Influence*; Robert J. Art and Kenneth N. Waltz, ed., *The Use of Force: Military Power and International Politics* (New York: Rowman & Littlefield, 2009).

## Revenge or Negative Reciprocity?

Models drawn from evolutionary psychology emphasize the centrality of environmental triggers and explain how contextual cues, such as those embedded in the social or institutional environment, activate different psychological strategies according to an ancestrally adaptive logic. When modern situations mirror these cues, the relevant psychological mechanisms become activated and shape our perceptions of threat and opportunity, and they instigate a repertoire of behavioral responses. Different environmental circumstances trigger different types of retaliatory aggression. The many forms of retaliatory aggression, such as revenge, can each be described along many dimensions, such as the *magnitude* of retaliation, the emotional *motivation* of retaliation, and the *function* of retaliation.

Even the speed of retaliation on its own can signal the underlying intention and meaning of behavior. In criminal law, this shows up in reduced sentences for “crimes of passion,” or those ostensibly committed in the heat of the moment, in a fit of anger or jealousy: Individuals acting under such duress usually receive a lesser sentence than those who engage in cold, calculated criminal planning. This can also play out at the international level. If, for example, Austria-Hungary had invaded Serbia immediately after the assassination of Archduke Franz Ferdinand in 1914, the Triple Entente would likely have accepted this response as legitimate, perhaps justified by public outrage in the wake of the unexpected murder. But Austria waited six weeks to launch an attack, and its actions instead appeared to be more of a calculated political effort to change the balance of power in Europe than a justified, if unfortunate, reaction to transgression.<sup>38</sup> Even across many traditional societies that have strong social norms against interpersonal violence, there is an understanding that revenge is inevitable and justified in certain circumstances.<sup>39</sup> One of the most pronounced and systematic illustrations relates to the relationship between female subordination and political order. Such tendencies predominate in clan-based governance

structures.<sup>40</sup> Examples of such patterns include phenomena such as honor killings, female genital mutilation, sex trafficking, and rape. To be clear, norms enforcing female subordination are not restricted to clan based governance structures. Indeed, such tendencies were also common in the American era of the Wild West. Rather, they provide examples of how strong social norms can often over-ride the institutional rule of law in a variety of circumstances, and such proclivities are particularly strong in the sexual arena. In other words, human nature engages in retaliatory violence in particular contexts, and audiences are quick to draw predictable inferences from the nature and context of these behaviors.

We can use these dimensions to identify how revenge is distinct from other forms of retaliation. For example, *negative reciprocity* is typically proportional to the initial harm, triggered by anger, and is aimed at recalibrating enemy preferences. In contrast, *revenge* is disproportional to the initial harm, often triggered by hatred, and functions to inflict harm on the enemy for the sheer pleasure of extracting vengeance (as well as possibly eliminating the adversary’s ability to deliver future harm). This can have the effect of establishing deterrence *ex ante*. We explore each in turn just below. Of course, issues of perception can come into play, and both sides may see the same situation in different ways. Third-party observers can play a role here by siding with the aggressor or instigator in a conflict.

### **Negative Reciprocity**

The most basic form of retaliation is tit-for-tat punishment, in which a harm received is responded to with a harm of relatively equivalent magnitude.<sup>41</sup> This type of retaliation is sometimes treated as synonymous with “punishment” but is more precisely understood as “negative reciprocity.”<sup>42</sup> In their application of negative reciprocity to international relations, Eder et al.<sup>43</sup> describe a Negative Reciprocity Norm as “involving a unitary set of beliefs favoring retaliation as the correct and proper way to respond to unfavorable treatment.” Punishment evolved as a strategy to make others pay for harms they inflicted. If such a strategy had

38 Jay Winter, ed., *The Cambridge History of the First World War* (Cambridge: Cambridge University Press, 2014).

39 Boehm, “Retaliatory Violence in Human Prehistory.”

40 Valerie M. Hudson, Donna Lee Bowen, and Perpetua Lynne Nielsen, “Clan Governance and State Stability: The Relationship Between Female Subordination and Political Order,” *American Political Science Review* 109, no. 3 (2015): 535-555.

41 Robert Axelrod, “An Evolutionary Approach to Norms,” *American Political Science Review* 80, no. 4 (December 1986); Axelrod, *The Evolution of Cooperation*.

42 Tim H. Clutton-Brock and Geoffrey A. Parker, “Punishment in Animal Societies,” *Nature* 373, no. 6511 (January 1995).

43 Paul Eder et al., “Punishing Those Responsible for the Prison Abuses at Abu Ghraib: The Influence of the Negative Reciprocity Norm (NRN),” *Political Psychology* 27, no. 6 (December 2006): 810.

not developed, then bullies would always have been able to win, and the kind of cooperation that allows complex society to develop would not have been possible. The prospect of punishment helps to deter future exploitation or criminal activity. It can also help salvage the possibility for future cooperation if the aggressor comes into line.<sup>44</sup> A great deal of work suggests that forms of so-called altruistic punishment, such as third-party punishment, evolved precisely to facilitate cooperation over time.<sup>45</sup> Somewhat counterintuitively, many forms of retaliation are socially productive in that they can improve bargaining and make compromise possible.<sup>46</sup>

## Desire and instincts toward revenge can take over, as satisfaction in the face of retribution comes to feed on itself.

A defining feature of negative reciprocity is that it facilitates cooperative bargaining with the delivery of measured punitive responses designed to signal information about interests and values. In other words, acts of punishment let the target know they have not sufficiently incorporated the attacker's interests into their calculations or are in violation of an agreement. Negative reciprocity must be measured and relatively proportional if cooperation is to be maintained or reestablished, which is a key departure from revenge. Sandra Bloom captures the reason for the measured (i.e., proportional) nature of negative reciprocity: "The injury and response must be balanced. An over-retaliatory response

provokes escalation while an under-retaliatory response provokes exploitation."<sup>47</sup> Accordingly, people are more accepting of retaliation when it is viewed as "symmetric" (e.g., "poetic" justice).<sup>48</sup>

Negative reciprocity predicts that punishment allows for the possibility of future cooperation as well as possibilities for forgiveness. Extant research on anger and punishment suggests that when the target of retaliation acknowledges and understands that a wrong has been done, punishers feel satisfied, anger is reduced, and forgiveness and reconciliation become possible.<sup>49</sup> If such acknowledgement does not occur, then escalation can result, particularly if escape is not an option. For example, the hostility of the People's Republic of China and South Korea toward Japan is fueled in part by Japan's perceived lack of remorse for past harms it perpetuated in those countries before and during World War II.<sup>50</sup> By contrast, Germany has accepted a different degree of responsibility for its war crimes and offered some reparations.<sup>51</sup> This does not necessarily result in forgiveness, but it does open the possibility for cooperation in the future.

### Revenge

In many ways, revenge can be understood as a more expressive emotional expression than the kind of tit-for-tat negative reciprocity, which is often more instrumental in nature. This is because negative reciprocity strives to change the opponent's behavior whereas revenge often only seeks the utter annihilation of the adversary. In contrast to negative reciprocity, which is motivated by anger and holds the possibility of reconciliation and future cooperation, revenge is the emotionally mediated psychological motivation or desire to harm for its own sake, expressing a form of hatred. Revenge attacks are more likely to be disproportionate, serving the evolutionary function of eliminating or reducing the target's ability to deliver future

44 Ben Seymour, Tania Singer, and Ray Dolan, "The Neurobiology of Punishment," *Nature Reviews Neuroscience* 8, no. 4 (April 2007); Michael E. Price, Leda Cosmides, and John Tooby, "Punitive Sentiment as an Anti-Free Rider Psychological Device," *Evolution and Human Behavior* 23, no. 3 (May 2002); Napoleon A. Chagnon, "Life Histories, Blood Revenge, and Warfare in a Tribal Population," *Science* 239, no. 4843 (February 1988).

45 Ernst Fehr and Simon Gächter, "Fairness and Retaliation: The Economics of Reciprocity," *The Journal of Economic Perspectives* 14, no. 3 (Summer 2000); Dominique J. F. de Quervain et al., "The Neural Basis of Altruistic Punishment," *Science* 305, no. 5688 (August 2004); James H. Fowler, "Altruistic Punishment and the Origin of Cooperation," *Proceedings of the National Academy of the Sciences* 102, no. 19 (May 2005).

46 Sandra L. Bloom, "Commentary: Reflections on the Desire for Revenge," *Journal of Emotional Abuse* 2, no. 4 (2001).

47 *Ibid.*

48 Thomas M. Tripp, Robert J. Bies, and Karl Aquino, "Poetic Justice or Petty Jealousy? The Aesthetics of Revenge," *Organizational Behavior and Human Decision Processes* 89, no. 1 (September 2002).

49 Mario Gollwitzer, Milena Meder, and Manfred Schmitt, "What Gives Victims Satisfaction When They Seek Revenge?" *European Journal of Social Psychology* 41, no. 3 (November 2010); McCullough, Kurzban, and Tabak, "Cognitive Systems for Revenge and Forgiveness"; Jeni L. Burnette et al., "Forgiveness Results From Integrating Information About Relationship Value and Exploitation Risk," *Personality and Social Psychology Bulletin* 38, no. 3 (November 2011); Todd H. Hall, *Emotional Diplomacy: Official Emotion on the International Stage* (Cornell: Cornell University Press, 2015).

50 John W. Dower, *Embracing Defeat: Japan in the Wake of World War II* (New York: W.W. Norton & Co., 2000).

51 Ronald W. Zweig, *German Reparations and the Jewish World: A History of the Claims Conference* (London: Frank Cass Publishers, 1987).

harm,<sup>52</sup> but the conscious motivation is simply the expected and intrinsic satisfaction of rebalancing the scales of experienced pain. People want to hurt others who have harmed them, even when they know that may not right the initial wrong done. It feels good, and people need not understand why. The instinct is automatic, effortless, and natural. While anger-fueled negative reciprocity yields psychological rewards contingent upon the target's understanding, hate-fueled revenge yields rewards that are contingent upon the target's suffering. That suffering, not just the target's understanding of the reasons for its suffering, is what satisfies the thirst for revenge.

Revenge most often occurs with no prior history of negative reciprocity or any attempt at reconciliation.<sup>53</sup> Men who lose comrades in combat are less inclined to negotiate with the other side and more inclined to do everything in their power to kill the enemy and extract vengeance for the lost brother. The motive can also grow out of failed attempts to compromise. Anyone who has ever watched people on the opposite side of the aisle become both more entrenched and more extreme as efforts to compromise fail has witnessed an example of this phenomenon. When bargaining through negative reciprocity fails, neither party may be willing to adjust its preferences in ways that would allow cooperation to be reestablished. This psychological stalemate can then incentivize revenge over negative reciprocity: When the adjustment of a target's preferences is perceived as increasingly unlikely or impossible, the angry individual or group must choose whether to accept the new state of affairs (e.g., through avoidance or submission) or to escalate the conflict. In these instances, an adversary's very existence may come to represent an existential threat: The challenge then shifts from restructuring the adversary's preferences through anger and negative reciprocity to weakening or eliminating the adversary altogether to reduce or eliminate the damage it can impose. The goal is then no longer to persuade and coerce but to enervate and eliminate, through brute force.

Importantly, a focus on revenge as only a consequence of the *imposition of a harm* misses at least half of the picture: Evolutionarily, the *withholding of a benefit*, is the functional equivalent

of the imposition of a cost; revenge can be triggered not only in response to harms delivered but also in response to benefits withheld. Benefits can take many forms, such as trade deals over items the recipient considers essential and cooperation on international issues. They can also relate to status and prestige concerns, as when one side refuses to give public acknowledgement and status the other side believes it deserves. This approach suggests that many cases of what appear to be "preemptive attacks" may be acts of revenge triggered by the subjective perception of status benefits having been systematically and, to one side, unjustifiably withheld.

Denial of benefits can lead to violence even in the absence of obvious direct or immediate provocation. Germany's instigation of World War II provides perhaps the iconic illustration of a war that was domestically sold as both a justified attempt to rectify the disproportionate material harms and the withholding of status benefits imposed on the German population by the Allies after World War I. Hitler's ability to characterize an event driven by his personal ambitions as a service rendered to the German people to regain their proper status in the world was arguably one of the most critical factors in Germany's provocation.<sup>54</sup> Japan's involvement in the war, subsequent to its conquest of Manchuria, similarly can be seen in light of its frustration at not receiving the international status and deference Tokyo felt it deserved.<sup>55</sup>

*Negative reciprocity* deters by adjusting adversary preferences directly through anger and proportional forms of punishment, holding open the possibility of post-conflict cooperation or reconciliation as incentive. On the other hand, *revenge* is fueled by hatred and deters through the imposition of disproportionate and often maximum harm, including extermination, to render counter-retaliation unlikely or impossible.

## The Mechanisms of Revenge

As we have suggested, one of the primary reasons deterrence can be so effective, despite the irrationality of a nuclear second strike, is that the threatened retaliation is motivated by powerful emotions that trigger physiological rewards. Recent

52 McCullough, Kurzban, and Tabak, "Cognitive Systems for Revenge and Forgiveness"; Sell, Tooby, and Cosmides, "Formidability and the Logic of Human Anger."

53 Corinna Carmen Gayer, Shiri Landman, Eran Halperin, and Daniel Bar-Tal, "Overcoming Psychological Barriers to Peaceful Conflict Resolution: The Role of Arguments about Losses," *Journal of Conflict Resolution* 53, no. 6 (2009): 951-975.

54 Ian Kershaw, *Hitler: 1889-1936 Hubris* (New York: W.W. Norton & Co., 2000); Paul Fussell, *Wartime: Understanding and Behavior in the Second World War* (Oxford: Oxford University Press, 1989); John Keegan, *The Second World War* (UK: Random House, 2011).

55 Ian Kershaw, *Fateful Choices: Ten Decisions That Changed the World, 1940-1941* (New York: Penguin, 2007).

evidence, particularly in conflict scenarios, reveals an important distinction between two emotions that are central to group conflict generally and to deterrence specifically. These emotions — anger and hate — provide the motivational foundations for the systems of negative reciprocity and revenge described above.

### Anger

Anger is key to the operation and recognition of negative reciprocity. It is a functional component of a complex motivational system in humans designed to resolve conflicts of interest in favor of the angry individual. Anger sends the signal to another that the attacker's welfare has been undervalued; its function is to adjust or "recalibrate" the preferences of another such that they are encouraged to place greater weight on the angry individual's welfare.<sup>56</sup> In strategic terms, the goal would be trying to change the opponent's behavior by altering their cost-benefit analysis in the attacker's favor. In this sense, anger is the equivalent of the latent threat of force. Its purpose is well illustrated by Schelling,<sup>57</sup> who noted: "The threat of pain tries to structure someone's motives, while brute force tries to overcome his strength." In our approach, anger is the automatic psychological mechanism that operates to re-structure someone's motives (i.e., recalibrate their preferences) through the imposition of costs (i.e., threat of violence) or the withdrawal of benefits (i.e., threat of suspending future cooperation). In this way, emotion operates strategically in the context of negative reciprocity by functioning to shift an adversary's preference structure in the opponent's favor.

### Hatred

Anger may fuel responses toward tractable enemies as opponents seek to force their adversary to recalibrate with the hope of future cooperation, but in the face of implacable enemies, hatred can spontaneously erupt. Desire and instincts toward revenge can take over, as satisfaction in the face of retribution comes to feed on itself. In this way, the goal of revenge may move beyond a means-to-an-end process, as the feeling provides enough motivation and reinforcement to generate revenge-seeking behavior. A drive to seek revenge in these situations becomes motivated by the feeling that

emerges through hormones such as testosterone and adrenaline. The hormonal regulation of a feeling of pleasure associated with retaliation is unsurprising in an evolutionary context; it is sufficient for natural selection to shape nervous systems with a "desire" or "taste" for revenge when adaptively appropriate. This is similar to the human desire for sex or high-caloric food for their own sakes, rather than because people consciously expect them to maximize their reproductive success. Babies do not need to know that more calories helped improve their ancestors' odds of survival to prefer sugar water to plain water. After all, the system works best, and most efficiently, when it relies on reinforcement mechanisms that do not require rational deliberation or attention to operate effectively. This is why preferences become automatic and effortless. Evolutionary processes seek to maximize the chance for reproductive success of the most useful variations in human traits. This is because even tiny advantages aggregate over time. Evolution is neurocomputational: It operates over billions of people across millions of years. So even if something looks counterproductive in some cases, it can prove successful over long periods of time if it results in reproductive fitness advantages of even tiny proportions for close kin.

How do such mechanisms operate? Neurological evidence has shown that when subjects are made to consider the prospect of retaliation, reward centers in the brain are flooded with activity, releasing powerful endogenous motivators for such retaliation.<sup>58</sup> Despite the expectation of retaliatory catharsis, however, some studies have found that the majority of subjects end up feeling worse after having inflicted retaliation.<sup>59</sup> A few things help explain this: The motivation for action sometimes differs from the experience of the consequences of that action. Retaliation can serve more than one purpose, even if it is motivated by a single drive. The consequence of inflicting retaliation may feel bad precisely because it signals the failure to convince the other to change its behavior.

The delicate psychological balance between negative reciprocity driven by anger and revenge fueled by hatred helps to explain the challenge of war termination and the avenues for peace following wars, genocides, civil conflict, or other kinds of institutionalized discrimination such as

56 Sell, Tooby, and Cosmides, "Formidability and the Logic of Human Anger."

57 Schelling, *Arms and Influence* (italics added).

58 De Quervain et al., "The Neural Basis of Altruistic Punishment."

59 Kevin M. Carlsmith, Timothy D. Wilson, and Daniel T. Gilbert, "The Paradoxical Consequences of Revenge," *Journal of Personality and Social Psychology* 95, no. 6 (2008); Ulrich Orth, "Does Perpetrator Punishment Satisfy Victims' Feelings of Revenge?" *Aggressive Behavior* 30, no. 1 (January 2004); Brad J. Bushman, "Does Venting Anger Feed or Extinguish the Flame? Catharsis, Rumination, Distraction, Anger, and Aggressive Responding," *Personality and Social Psychology Bulletin* 28, no. 6 (June 2002).

apartheid. It is easy to see how anger can slide into hatred and tractable opponents can become intractable enemies. Halperin et al.<sup>60</sup> show that, in the context of the Arab-Israeli conflict, inducing anger in subjects through an experimental mood-manipulation technique increased compromise-seeking in negotiations while inducing hatred in subjects reduced their support for compromise. In this example, hatred is triggered, in part, by the perception of an out-group's "inability to undergo positive change."<sup>61</sup> Put another way, hatred is a response to the perception that the target's preference structure cannot be recalibrated and that one cannot avoid future costs inflicted by the target.

What is significant is not that hatred makes compromise intolerable but that anger makes compromise *possible*. Similarly, in their study of vicarious retribution, Lickel et al.<sup>62</sup> note that the success of the Truth and Reconciliation Commission in South Africa was premised upon the "need for understanding but not for vengeance, a need for reparation but not for retaliation." In effect, the commission was chartered with the task of manipulating a deescalatory slide from revenge to negative reciprocity, from hatred to anger, and ultimately toward reconciliation and understanding. Negative reciprocity motivates anger and retaliation while opening the possibility for reconciliation and understanding. It is only when anger slides into hatred that conflict escalates toward intractability.

### **Cueing Negative Reciprocity and Revenge**

Revenge operates to prevent future exploitation by promising retaliation that no one wants but that everyone believes will happen if certain violations

occur.<sup>63</sup> Different emotions trigger specific kinds of responses, just as different situations spark particular responses. The contrast between intra-war deterrence and classic nuclear deterrence provides such an illustration. The recognition and anticipation of these strategies (negative reciprocity versus revenge) is made possible by examining the environmental cues at play in each system.

Environmental and contextual cues can obviously serve as triggering mechanisms for aspects of motivation and behavior. Contexts of intra-war deterrence often trigger the operation of negative reciprocity, in which threats are needed to rebuff limited attacks without producing a full-scale conflagration. Almost all the literature on limited war, both empirical (such as that focused on Vietnam or Korea<sup>64</sup>) as well as theoretical<sup>65</sup> (often framed around prospects for winning a limited nuclear war) would come under this rubric of negative reciprocity. This literature has argued that for limited war to work, the response must be proportionate and function clearly to adjust the enemy's preferences. Although the distinction between changing the enemy's preferences and decreasing its capabilities is not always clear, the goal is to make the enemy change its behavior without engaging in an all-out war in which one side or the other risks decimation. This retains the possibility of engaging in cooperative behavior in the future.

By contrast, situations such as those outlined in classic nuclear-deterrence theory trigger the operation of revenge, in which the credibility of the threat to retaliate is designed to prevent a first strike. Classical deterrence theorists argued that if two adversaries could credibly commit to deliver unacceptably costly retaliation in response to the other's first strike, deterrence between them would hold.<sup>66</sup> As discussed earlier, scholars noted that

---

60 Eran Halperin et al., "Anger, Hatred, and the Quest for Peace: Anger Can Be Constructive in the Absence of Hatred," *Journal of Conflict Resolution* 55, no. 2 (April 2011).

61 Ibid.

62 Brian Lickel et al., "Vicarious Retribution: The Role of Collective Blame in Intergroup Aggression," *Personality and Social Psychology Review* 10, no. 4 (November 2006).

63 Boehm, "Retaliatory Violence in Human Prehistory"; McCullough, Kurzban, and Tabak, "Cognitive Systems for Revenge and Forgiveness"; Christopher Boehm, *Blood Revenge: The Enactment and Management of Conflict in Montenegro and Other Tribal Societies* (Philadelphia: University of Pennsylvania Press, 1984).

64 Stephen Peter Rosen, "Vietnam and the American Theory of Limited War," *International Security* 7, no. 2 (October 1982); Yuen Foong Khong, *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965* (Princeton: Princeton University Press, 1992).

65 Colin S. Gray and Keith Payne, "Victory Is Possible," *Foreign Policy*, no. 39 (Summer 1980); Desmond Ball, *Can Nuclear War Be Controlled?* (London: International Institute for Strategic Studies, 1981); Sidney D. Drell and Frank Von Hippel, "Limited Nuclear War," *Scientific American* 235 (November 1976); Thomas C. Schelling, "Bargaining, Communication, and Limited War," *Conflict Resolution* 1, no. 1 (1957).

66 Brodie, *Strategy in the Missile Age*; Snyder, *Deterrence and Defense: Toward a Theory of National Security*; Schelling, *The Strategy of Conflict*.

this required rational actors to commit to a course that was manifestly irrational.<sup>67</sup> Second-strike deterrence therefore faces a glaring problem: Once a country has sustained catastrophic damage, an all-out retaliatory strike can neither “win” the war nor limit the damage already experienced. In other words, there is no rational basis for retaliation once one’s fate is sealed. Yet such retaliation is precisely what deterrence requires and what human psychology delivers.

The psychology of revenge clearly reconciles this problem of making a futile second strike credible and allows second-strike deterrence to hold where classical economics suggests it should not.<sup>68</sup> Notably, a number of strategies have been designed and implemented to make this threat more credible from the perspective of economic rationality. Threats of retaliation can be made more credible by pre-delegation to commanders in the field, who may be less than controllable in the heat of battle. In addition, standard operating procedures can be implemented to ensure survivability, such as by invoking launch once it appears an adversary’s weapons are airborne to prevent the loss of one’s own missiles. Moreover, attempts to implement automaticity systems, such as the Soviet “Dead Hand” system, can also increase the credibility of threatened response.<sup>69</sup> Such efforts, while eminently sensible from a rationalist perspective, appear redundant and unnecessary from a psychological standpoint. Although second-strike deterrence represents a logical inconsistency for rational actors, few doubt that such retaliation would occur. Policymakers and lay people alike recognize that retaliation would be forthcoming not because they recognize the theoretical requirements of deterrence but because all humans instinctively recognize situations in which revenge is not only likely but also emotionally inevitable even when logically futile. Our shared psychology enables people to quickly and reliably recognize the power and pleasure of retaliation and revenge in the face of such an attack regardless of its logical value from a material standpoint. While irrationality may undercut formal notions of credibility for second-strike retaliation, the universal human

understanding of and drive for revenge more than compensates. Few doubt that retaliation would ensue in the face of assault, even if such reaction offers little prospect of victory or salvation. Even in a court of law, crimes of passion receive significant mitigation in sentencing precisely because everyone believes individuals are less responsible for their actions under such circumstances.

In short, negative reciprocity and revenge operate according to distinct logics. This logic is reflected in and triggered by distinct emotional triggers as well as distinct contextual cues that manifest within situations of intra-war versus second-strike deterrence. Deterrence is fundamentally about the appropriate use of retaliatory threats to affect adversary behavior. The distinction between negative reciprocity and revenge, as well as an awareness of their distinct emotional antecedents and context-specificity, can deepen understanding of how, when, and why deterrence works or fails in particular situations.

### Dynamics of Revenge, Within and Between

Having described the implications of revenge for deterrence and distinguished it from other forms of retaliation as illustrated by the contextual cues that trigger these mechanisms, it is useful to examine ways in which this psychology is expressed within and between individuals, groups, and states.

One of the more apparent characteristics of revenge at the group level is that avengers tend to target out-group members indiscriminately. Apropos of this recognition, one common defining attribute of “weapons of mass destruction” and modern terrorism generally is that they are relatively indiscriminate. The anthropologist Raymond Kelly has referred to this as “social substitution,” which occurs when individuals hold all or any out-group members responsible for the transgressions of one member, in effect treating the group as a unitary actor.<sup>70</sup> This is referred to as “third-party revenge” or “vicarious retribution.”<sup>71</sup> Osama bin Laden’s fatwas and similar indictments that hold all

67 Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence* (Cambridge: Cambridge University Press, 2000); Peter R. Lavoy, Scott D. Sagan, and James J. Wirtz, ed. *Planning the Unthinkable: How New Powers Will Use Nuclear, Biological, and Chemical Weapons* (Cornell: Cornell University Press, 2000); Powell, *Nuclear Deterrence Theory: The Search for Credibility*; Robert S. McNamara, “The Military Role of Nuclear Weapons: Perceptions and Misperceptions,” *Foreign Affairs* (Fall 1983); Robert Powell, “Nuclear Deterrence and the Strategy of Limited Retaliation,” *American Political Science Review* 83, no. 2 (June 1989); Jervis, “Deterrence Theory Revisited.”

68 John J. Mearsheimer, “Nuclear Weapons and Deterrence in Europe,” *International Security* 9, no. 3 (Winter 1984-85); Keith B. Payne, *Deterrence in the Second Nuclear Age* (Lexington: University Press of Kentucky, 2015).

69 David E. Hoffman, *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy* (New York: Anchor, 2009).

70 Raymond C. Kelly, *Warless Societies and the Origin of War* (Ann Arbor: University of Michigan Press, 2000).

71 Lickel et al., “Vicarious Retribution: The Role of Collective Blame in Intergroup Aggression.”

Americans responsible for particular behavior can be seen in this category.<sup>72</sup> According to Michener,<sup>73</sup> a single attack from an out-group is often sufficient for individuals to ascribe an “enemy” image to that out-group, which enables vicarious retribution through the simple mental algorithm “one did it/they all did it.” Certainly, the scapegoating of entire groups based on the actions of some subset of individuals, such as the anti-Western bias of Islamic jihadists, or some Westerners’ antipathy toward all Muslims, falls into this category of treating all members of a group as the same regardless of individual culpability for bad actions.

Our approach suggests that this representation does not capture the whole picture. On the one hand, it is often true that an attack against one’s group can precipitate vicarious retribution. Of course, there can be different types of response, as was discussed above in relation to anger, negative reciprocity, proportionate response, and prospects for limited war. However, across the entirety of human civilization, especially under circumstances characterized by hatred and revenge, total annihilation of the enemy was more often the rule than the exception. Certainly sometimes what the victors do to the vanquished can incorporate an element of strategic or instrumental coercion, as Sherman’s march through Atlanta at the end of the Civil War is often understood to incorporate.<sup>74</sup> Still, examples abound of laying waste to the enemy, from the sacking of Carthage through the siege of Stalingrad to more recent examples of ethnic cleansing.<sup>75</sup> Examples of the wholesale destruction undertaken by the Mongol invaders, Alexander the Great, and many others since suggest that while some population assimilation may have occurred, most often as local women were kidnapped, captured, and raped, annihilation was the more typical response to attack.<sup>76</sup> Vicarious retribution can be triggered not only by surprise attacks but also by humiliation or defeat. Consider as evidence Hitler’s declaration that “we do not pardon, we demand vengeance!”<sup>77</sup> In the context of intergroup hostilities, blind vicarious retribution has historically proven the norm.

On the other hand, not all attacks from other states necessarily precipitate vicarious retribution from the victim’s group. To be sure, very weak states simply may not have the capability to retaliate, although in a globalized nuclear-armed world, the ability to hurt is increasingly independent of one’s

## Examples abound of laying waste to the enemy, from the sacking of Carthage through the siege of Stalingrad to more recent examples of ethnic cleansing.

conventional military strength. The violation of sacred values, or the search for status, may compel even objectively weaker states and non-state actors to challenge and sometimes defeat conventionally superior powers.<sup>78</sup> More relevant to our approach is the recognition that in principle there should be contexts in which an attack from an out-group generates anger but not necessarily the kind of hatred that would normally be prompted by a devastating surprise attack on domestic territory or nationals. This can happen in at least two ways.

First, the assailant may be a member of an out-group who is politically aligned with some members of the in-group. As behavioral and physiological studies have demonstrated, intra-alliance bonds can mitigate the otherwise escalatory effects of victory in an inter-group context.<sup>79</sup> The reason is simple: The cost of losing or weakening the alliance may be greater than the benefit of intra-alliance conflict-escalating behaviors. In other words, coalitions that need to stick together to successfully face future threats will not benefit from remaining preoccupied with trivial rivalries within the group in the short

72 Peter L. Bergen, *The Osama bin Laden I Know: An Oral History of al Qaeda’s Leader* (New York: Simon and Schuster, 2006).

73 Willa Michener, “The Individual Psychology of Group Hate,” *Journal of Hate Studies* 10, no. 1 (January 2012).

74 Shelby Foote, *The Civil War: A Narrative* (New York: Random House, 2011).

75 Norman M. Naimark, *Fires of Hatred: Ethnic Cleansing in Twentieth-Century Europe* (Cambridge: Harvard University Press, 2002).

76 John Bagnell Bury, *A History of Greece to the Death of Alexander the Great* (London: Macmillan and Co., Ltd., 1906).

77 Adam Hochschild, *To End All Wars: A Story of Loyalty and Rebellion, 1914-1918 Reprint edition* (Boston: Mariner Books, 2012).

78 Scott Atran and Jeremy Ginges, “Religious and Sacred Imperatives in Human Conflict,” *Science* 336, no. 6083 (May 2012); Ivan Arreguin-Toft, “How the Weak Win Wars: A Theory of Asymmetric Conflict,” *International Security* 26, no. 1 (Summer 2001).

79 Mark V. Flinn, Davide Ponzi, and Michael P. Muehlenbein, “Hormonal Mechanisms for Regulation of Aggression in Human Coalitions,” *Human Nature* 23, no. 1 (March 2012).

term. This is why conflicts characterized by negative reciprocity and cued by anger are more likely to exist in the context of intra-war rivalry. The 1837 Caroline affair between the United States and Britain regarding Canadian revolutionaries provides an example. British troops boarded a ship, the Caroline, sailed by members of the Canadian independence movement, led by William Mackenzie, and some American supporters. The British troops killed an American, burned the ship, and tossed it over Niagara Falls. This led Americans and Canadians to retaliate against the British; the British ship was destroyed as well. After several tit-for-tat attacks, the situation was resolved with the Ashburton-Webster Treaty of 1842, whereby Daniel Webster wrote that the necessity for preemptive self-defense must be characterized by “instant, overwhelming, and leaving no choice of means, and no moment for deliberation.” This “Caroline test” still provides the foundation for what has become a pillar of international law regarding the preemptive use of force.<sup>80</sup>

Not surprisingly, this tempering mechanism is reflected in humans’ biological and psychological architecture. Even though victory in competition tends to lead to higher testosterone levels and dominance displays among the victors across numerous contexts,<sup>81</sup> this reaction is muted when the defeated party is considered part of the in-group.<sup>82</sup> Aside from the obviously stabilizing effect this has on intra-group relations, this muted testosterone response mitigates the prospect of anger turning into hatred and, therefore, helps to mitigate the perceived need among the defeated for revenge against in-group members.

A second context in which an out-group attack may elicit anger but not hatred is the case in which the out-group is perceived to be internally fractured. This would support the strategic adage to “divide and conquer.” In this way, prospects for converting some out-group members into allies reduce the need, and desire, for total annihilation

of the out-group. If future cooperation is desirable, beneficial, and possible, it behooves antagonists to try to resolve a tractable conflict rather than turn an opponent into an intractable enemy. In contrast, when out-groups appear to operate relatively cohesively, there is evidence that individuals are more likely to hold groups collectively responsible for attacks of their individual members.<sup>83</sup> According

**Although leaders may find it challenging, in certain contexts, to restrain a public bent upon vengeful fervor, leaders may conversely find that lighting the wick of outrage is no simple matter either.**

to Lickel<sup>84</sup>: “If the group is perceived to be highly unified, then other members of that group are more likely to be blamed and targeted for retribution for the provocative acts of an individual group member.”

A good example of this dynamic exists in U.S.-Iranian relations. There is an emotional divide within the United States between those who see Iran as an implacable enemy, based largely on events surrounding the Iranian hostage crisis from 1979, and those who view Iran as America’s most logical ally in the region, given the history of close relations before 1979 and the fact that the two countries share a natural enemy in the self-proclaimed Islamic State, among other overlapping interests. Those who express a sense of betrayal and anger at Iran for overthrowing the shah and taking American personnel hostage also tend to place less emphasis on the U.S.-led coup against Iran’s democratically elected prime minister, Mohammad Mossadegh, in 1953, when the United States helped reinstate the shah, who was friendlier

80 Robert Y. Jennings, “The Caroline and McLeod Cases,” *The American Journal of International Law* 32, no. 1 (January 1938): 82-99; Martin A. Rogoff and Edward Collins, Jr., “The Caroline Incident and the Development of International Law,” *Brook. Journal of International Law* 16 (1990): 493.

81 John Archer, “Testosterone and Human Aggression: An Evaluation of the Challenge Hypothesis,” *Neuroscience & Biobehavioral Reviews* 30, no. 3 (2006).

82 John D. Wagner, Mark V. Flinn, and Barry G. England, “Hormonal Response to Competition Among Male Coalitions,” *Evolution and Human Behavior* 23, no. 6 (November 2002).

83 Adam Waytz and Liane Young, “The Group-Member Mind Trade-Off Attributing Mind to Groups Versus Group Members,” *Psychological Science* 23, no. 1 (December 2011).

84 Lickel et al., “Vicarious Retribution: The Role of Collective Blame in Intergroup Aggression,” *Personality and Social Psychology Review* 10, no. 4 (2006): 372-390.

to U.S. and British interests.<sup>85</sup> They also tend to reject the possibility that any benefit could result from closer relations; indeed, these are the people who strongly reject the Iranian nuclear deal when most experts argue that the United States benefits more than it loses from the agreement. On the other hand, those who see Iran as more useful as a potential future ally than adversary tend to argue that closer relations may incentivize the Iranians to shift their behavior in ways that are more conducive to American interests.<sup>86</sup> Similar arguments are made by those who seek to incorporate North Korea into the international community, hoping that greater economic integration in particular will provide enough of an incentive to offer political leverage as well.<sup>87</sup> Importantly, these differences in the level of anger and resentment, clearly reflected in generational differences in attitudes toward Iran, reflect how distinct emotional perspectives can yield different inferences about why Iran acts as it does, and suggests divergent responses to such behavior.

These two points taken together suggest that the well-studied “enemy image” in international relations may obscure important nuances that deserve to be conceptually unpacked. Specifically, this question relates to the uncertainty that states and leaders confront in trying to figure out the nature of the enemy they confront: Can the opponent be enticed to cooperate to the benefit of both, or does the other side present an intractable enemy who should be fought sooner rather than later? For example, the “enemy” image characterizes a state as monolithic, evil, strategically opportunistic, and yielding only in the face of the perceiver’s commitment and strength.<sup>88</sup> Indeed, once the out-group is perceived as monolithic and evil (i.e., cohesive and not subject to preference recalibration), it is likely to inspire hatred and attacks that will precipitate escalatory and indiscriminately vicarious vengeance. But to the extent that the enemy is seen as a group that can be divided and conquered, or one responsive to coercion, negative reciprocity can offer a more

useful strategy because it retains the possibility of future cooperation.

It may be the case that inter-group relations are naturally characterized by a bias toward perceiving out-groups as evil and monolithic, particularly since the consequences of misjudgment are much graver for perceiving the enemy as friendly than the reverse.<sup>89</sup> Nevertheless, it would be folly to mistake mere hurdles for inevitabilities, especially given the natural abilities of humans to manipulate and widen the bonds of group membership.<sup>90</sup> These dynamics reveal the danger of misidentifying an enemy whose preferences can be recalibrated for one whose preferences cannot, which risks precipitating the very outcome each side may wish to prevent: escalation in retaliatory violence. In addition, the potential for future cooperation may be lost as well, which can entail an extremely high cost over time. Tragically, misidentifying the character of the opponent may be one important mechanism that fuels the recurrent security dilemma in international relations whereby actors make attributions about others’ intentions that are more hostile than is actually the case. This concern can easily trigger an escalatory dynamic as each side makes assumptions about the other that inspire increased weapons procurement for purposes of defense, which are simultaneously understood by the other side as signaling escalatory or hostile intent.<sup>91</sup>

Aside from perceptions of group cohesion, which appear to mediate the degree to which groups are prepared to engage in escalatory revenge, a second major pathway from individual psychology to coalitional dynamics lies in the role of leaders. Again, however, the natural tendency in leadership seems biased toward revenge over negative reciprocity in the face of an out-group attack. Leaders tend to be more prototypical on relevant in-group attributes and more susceptible to out-group threat.<sup>92</sup> By extension, leaders are more likely to seek revenge and be held as targets for revenge, given their group prototypicality and symbolism. Nevertheless, leaders are not merely powerless

85 Walter LaFeber, *American Age: U.S. Foreign Policy at Home and Abroad, From 1750 to the Present* (New York: W.W. Norton & Company, 2nd edition, 1994); Stephen Kinzer, *All the Shah's Men: An American Coup and the Roots of Middle East Terror* (Hoboken: John Wiley & Sons, 2003).

86 Stephen Kinzer, *Reset: Iran, Turkey, and America's Future* (New York: Macmillan, 2010).

87 Harry G. Broadman, "Time To Try An Economic Carrot Approach With North Korea?" *Forbes*, September 30, 2017, <https://www.forbes.com/sites/harrybroadman/2017/09/30/time-to-try-an-economic-carrot-approach-with-north-korea/#48232b5712bc>.

88 Keith L. Shimko, *Images and Arms Control: Perceptions of the Soviet Union in the Reagan Administration* (Ann Arbor: University of Michigan Press, 1992); Robert Jervis and Jack Snyder, *Dominoes and Bandwagons: Strategic Beliefs and Great Power Competition in the Eurasian Rimland* (New York: Oxford University Press, 1991).

89 Martie G. Haselton and David M. Buss, "Error Management Theory: A New Perspective on Biases in Cross-Sex Mind Reading," *Journal of Personality and Social Psychology* 78, no. 1 (January 2000).

90 Robert M. Sapolsky, "Social Status and Health in Humans and Other Animals," *Annual Review of Anthropology* 33 (October 2004).

91 Robert Jervis, *Perception and Misperception in International Relations* (Princeton: Princeton University Press, 1976).

92 Michael A. Hogg, "A Social Identity Theory of Leadership," *Personality and Social Psychology Review* 5, no. 3 (2001).

drones of the masses. Their ability to influence the direction of inter-group conflict is a function of their ability to frame events, sway opinion, and mobilize resources.

Depending on the context, in-group elites may wish to either moderate or facilitate vengeful responses to an out-group attack, whether for the sake of political expediency, strategic necessity, or alliance considerations. For example, President Grover Cleveland and his successor, William McKinley, each tried but ultimately failed to restrain a Congress and public that were increasingly eager to expel Spain from Cuba. It was the destruction of the U.S.S. *Maine* in 1898 that helped tip the political scales toward war with Spain. Although no direct evidence of Spanish culpability was unearthed, the relevant image that Americans had of Spain was that of an attacker, not an ally; there was therefore no room for pause or deliberation after the ship's destruction. As a counterfactual, one might expect that had Spain been a more valuable ally, American elites would have made more of an effort to excuse or interpret events in a different light. In reality, U.S. interests — both ideological and economic — lay with the Cubans. Perceived Cuban suffering at the hands of Spanish imperialists, and the volume of American trade and investment in Cuba being greater than even that with Spain, conspired to mute McKinley's ability to restrain his country's outrage at ostensibly hostile Spanish actions in the Caribbean. McKinley would later lament that "but for the inflamed state of public opinion, and the fact that Congress could no longer be held in check, a peaceful solution might have been held."<sup>93</sup>

Although leaders may find it challenging, in certain contexts, to restrain a public bent upon vengeful fervor, leaders may conversely find that lighting the wick of outrage is no simple matter either. This was Woodrow Wilson's conundrum in wanting to control the peace without entering World War I and, having opposed U.S. involvement, then having to rouse sentiment in its favor at a time of entrenched isolationist sentiment. This was also Franklin D. Roosevelt's challenge as he simultaneously sought to support the Allies in their struggle against Nazi Germany while reassuring an ambivalent public that the United States would not enter the war. According to Schuessler,<sup>94</sup> Roosevelt's dilemma was to reconcile two contrasting realities: 70 percent of Americans wanted to stay out of the war, but 70 percent wanted to see Hitler defeated at all costs. Despite the strategic threat posed by Germany, it

was the surprise attack at Pearl Harbor that would mobilize popular sentiment in such a way as to allow Roosevelt to more overtly support the Allies.

### **"Be This the Whetstone of Your Sword"**

Revenge is the product of complex psychological mechanisms that evolved in response to the adaptive problem of deterring adversaries in ancestral coalitional environments. Synthesizing research across disciplines, we expose the psychological underpinnings of revenge, its implications for the functioning of deterrence in its role in international relations, and its differentiation from other forms of retaliation such as negative reciprocity. Furthermore, we explore its emotional correlates and contextual triggers in order to illustrate the potency, plasticity, and application of these systems across seemingly disparate domains of international politics such as intra-war conflict and nuclear deterrence. Revenge occurs naturally and automatically; it is the psychological — even if latent — dynamic that makes deterrence possible.

Importantly, our distinction between revenge and negative reciprocity provides theoretical scope for understanding the nature of revenge and provides scholars with a useful typology for beginning to understand the many ways that states and non-state actors are likely to respond to threat. According to our typology, *negative reciprocity* operates to recalibrate adversary preferences through anger and the proportional delivery of punishment, and it holds out the possibility of post-conflict reconciliation. In contrast, *revenge*, which is motivated by more intense emotions such as hatred, operates to impose disproportionate and often maximum harm in order to render counter-retaliation unlikely or impossible. In the face of existential threat, revenge overwhelms the cost-benefit calculations that would otherwise lead rational actors to accept sunk costs and, instead, to return spiteful destruction on the attacker. This desire is endogenously motivated through the neuroendocrine system. Whereas negative reciprocity operates to resolve conflict via preference restructuring and bargaining, revenge resolves conflict by crippling or eliminating the adversary. Although it may seem odd to think of revenge as a "conflict resolution" mechanism, that is its proper domain. As Daly and Wilson<sup>95</sup> aptly note: "Killing

93 Richard F. Hamilton, *President McKinley, War and Empire: President McKinley and the Coming of War, 1898* (Abingdon: Routledge, 2006).

94 John M. Schuessler, "The Deception Dividend: FDR's Undeclared War," *International Security* 34, no. 4 (Spring 2010): 153.

95 Daly and Wilson, *Homicide*.

one's antagonist is the ultimate conflict resolution technique." However, its effectiveness to end one conflict is often balanced by its ability to generate new conflicts. Disproportionate revenge has often led to friends, relatives, and others who shared a sense of community and identity with the target to respond in kind, unleashing a never-ending escalating spiral, as is often seen in civil conflicts, terrorism, and other forms of political violence. Famous family feuds often have this character as well, demonstrating that the psychological dynamics we explicate are not restricted to nuclear deterrence or nation-state behavior but, rather, that the behavior of leaders and elites reflects basic human psychological mechanisms operating within the context of the larger political domain. If the strategy of revenge was successful enough to provide even a small fitness advantage over time and across many people, especially through the annihilation of enemies, which would prevent consequent retaliation, the instinct could easily be preserved, at least among some percentage of the population, even if much of the time it may appear counterproductive in any single instance.

This perspective allows scholars and policymakers to potentially infer internal motivations from behavioral responses, which, in turn, can help policymakers avoid costly errors such as falsely identifying an adversary as implacable when bargaining is more possible than it might outwardly seem. Policymakers can also potentially distinguish between these types of adversaries by the contextual environments in which they occur, as well as the emotions they elicit. For example, anger-fueled negative reciprocity tends to occur in environments characterized by intra-war conflict, while revenge, fueled by hatred, emerges in the case of all-out war, including the prospect of nuclear conflagration. Their emotional manifestations provide the motivating force that both signals and sustains their respective functions. These predictive inferences and implications are the necessary next steps for researchers as we continue to explore the evolved psychology of threat perception in coalitional contexts.

An evolutionary perspective also reveals that revenge may be sought not only by the direct imposition of costs and harm but also by the systematic withholding of benefits that affect a group's status or resources. When revenge is triggered by the withholding of benefits, it is no less "retaliatory" simply because it is, behaviorally

speaking, a first strike. Many cases of attacks that appear preemptive in nature are better explained as acts of revenge triggered by the subjective perception of the systematic withholding of status or other material benefits. In other words, the perception of injustice can lead to anger and eventually hateful vengeance even in the apparent absence of obvious external or physical provocation. This is applicable in the case of rising powers, for example, which are likely to experience the widening differential between their material power and relative status as humiliating, precipitating a slide from angry contempt to hateful spite toward those that distribute rights and benefits in the international system. This can also happen when material benefits are withheld.

Theoretically, evolutionary models provide novel information that extends models based on traditional notions of rationality by moving beyond purely cognitive readings of credibility and deterrence to offer insight into how specific emotions or environmental contexts can serve as motivating cues for behavioral responses. Importantly, evolutionary models are not constructed along the lines of traditional economic definitions of "rationality," such as those based on immediate cost-benefit calculations. Rather, they are formulated according to an organism's long-term reproductive success, which includes emotional short cuts and cognitive heuristics that, while not appearing rational from a classical economic perspective, serve a much deeper rationality designed to facilitate the survival of the organism.

Many models of rationality assume that individual rationality, defined in classical economic terms, can lead to collective rationality.<sup>96</sup> Others acknowledge that collective irrationality or sub-optimal outcomes can paradoxically result from individuals pursuing their rational self-interest.<sup>97</sup> Yet neither perspective has paid much attention to the ultimate origins or ecological validity of such preference structures. A major contribution of evolutionary models is that they offer a means to better understand human decision making and preference structures within given contexts. Whether a behavior is rational from an economic perspective reflects only one real but very limited and narrow facet of rationality. Evolutionary psychology interrogates the adaptively functional structure of decision-making systems and preferences, and the environmental triggers that activate a wide variety of systems. In contrast, rational choice models often assume a set

96 Schelling, *The Strategy of Conflict*.

97 Russell Hardin, "Collective Action as an Agreeable N-Prisoners' Dilemma," *Behavioral Science* 16, no. 5 (September 1971); Mancur Olson, Jr., *The Logic of Collective Action: Public Goods and the Theory of Groups*, Rev. Ed (New York: Schocken, 1971).

of preferences *a priori* and examine the effects of environmental constraints on those preferences. This constitutes an important point of theoretical divergence in these models. This does not mean such models must necessarily exist in opposition. Rather, evolutionary models can inform the lacuna that exists in rational models that fail to identify the origin of preferences; preferences easily emerge from the logic of reproductive success from an evolutionary perspective. In this light, a full appreciation of the purpose and function of revenge offers a universal basis for the emotional motivations that undergird, however implicitly, more economically rational notions of deterrence.

The causes of war are well studied in international politics, particularly from a more traditional rationalist perspective. Increasingly, scholars have turned to the behavioral sciences and research on human emotion to complement understanding of war and to deepen society's understanding of the triggers of political conflict. The psychological investigation of revenge is critical for international relations scholarship because revenge has for centuries remained among the most common motivations for hostility, and it involves the operation and expression of a very intense set of human emotions. Revenge is a notoriously stubborn, recurrent, and tragically prevalent motivation for political violence at all levels of social organization, and though the complexity and evolutionary novelty of international political structures may modify the behavioral expression of this basic human tendency, revenge remains a disturbingly central feature of international, sub-state, and personal conflicts. This recognition alone warrants a deeper appreciation of its significance in calibrating human conflict and supporting the structure of deterrence. 

**Rose McDermott** is the David and Mariana Fisher University Professor of International Relations at Brown University and a Fellow in the American Academy of Arts and Sciences. She received her Ph.D. (Political Science) and M.A. (Experimental Social Psychology) from Stanford University and has taught at Cornell, University of California-Santa Barbara, and Harvard. She has held fellowships at the Radcliffe Institute for Advanced Study, the Olin Institute for Strategic Studies, and the Women and Public Policy Program, all at Harvard University. She has been a fellow at the Stanford Center for Advanced Studies in the Behavioral Sciences twice. She is the author of four books, a co-editor of two additional volumes, and author of over two hundred academic articles across a wide variety of disciplines encompassing topics such as experimentation, emotion and decision making, and the biological and genetic bases of political behavior.

**Peter K. Hatemi** is Distinguished Professor of Political Science, Microbiology, and Biochemistry at The Pennsylvania State University. He was trained in genetic epidemiology at the Queensland Institute of Medical Research (QIMR). He continued his post-doctoral study in Human Genetics, Psychology, and Psychiatry at the Virginia Institute for Psychiatric and Behavioral Genetics in the Medical College of Virginia. He is primarily interested in advancing the study of the neurobiological mechanisms of social and political behaviors and utilizing advanced methods in genetics, physiology, endocrinology, and neurology in order to better understand human decision making and preferences in complex and dynamic political environments. He is also an active member of the Institut for Statskundskab at Syddansk Universitet, VIPBG and the genetic epidemiology lab at QIMR.

Dr. Hatemi's recent work has appeared in the American Journal of Political Science, Behavior Genetics, Demography, Evolution and Human Behavior, Journal of Politics, Political Psychology, Science, and Social Forces among other venues. His book, co-edited with Rose McDermott, *Man is by Nature a Political Animal* includes applications of evolution, genetics, primatology, neuroscience, and physiology, to understand political preferences.

**Anthony C. Lopez** received a Ph.D. from Brown University in Political Science and is Assistant Professor of International Relations and Political Psychology at Washington State University. His research investigates war as the product of an evolved coalitional psychology, and examines the relationship between inter-group conflict and intra-group cooperation from an adaptationist perspective.

