*Behavioral and Brain Sciences*

**cambridge.org/bbs**

## Target Article

**Cite this article:** De Dreu CKW, Gross J. (2019) Revisiting the form and function of conflict: Neurobiological, psychological, and cultural mechanisms for attack and defense within and between groups. *Behavioral and Brain Sciences* **42**, e116: 1–66. doi:10.1017/S0140525X18002170

Target Article Accepted: 13 September 2018
Target Article Manuscript Online: 25 September 2018
Commentaries Accepted: 11 February 2019

**Keywords:**
behavioral game theory; biobehavioral approach–avoidance; coevolution of prosociality and aggression; conflict; conflict intervention; cultural institutions; intergroup relations; psychological adaptations

**What is Open Peer Commentary?** What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 15) and an Author's Response (p. 45). See bbsonline.org for more information.

© Cambridge University Press 2019

CAMBRIDGE UNIVERSITY PRESS

# Revisiting the form and function of conflict: Neurobiological, psychological, and cultural mechanisms for attack and defense within and between groups

Carsten K. W. De Dreu and Jörg Gross

Institute of Psychology, Leiden University, 2300 RB Leiden, The Netherlands Center for Research in Experimental Economics and Political Decision Making (CREED), University of Amsterdam, 1018 WB Amsterdam, The Netherlands and Institute of Psychology, Leiden University, 2300 RB Leiden, The Netherlands
c.k.w.de.dreu@fsw.leidenuniv.nl https://www.universiteitleiden.nl/en/staffmembers/carsten-de-dreu
mail@joerg-gross.net http://www.joerg-gross.net

**Abstract**

Conflict can profoundly affect individuals and their groups. Oftentimes, conflict involves a clash between one side seeking change and increased gains through victory and the other side defending the status quo and protecting against loss and defeat. However, theory and empirical research largely neglected these conflicts between attackers and defenders, and the strategic, social, and psychological consequences of attack and defense remain poorly understood. To fill this void, we model (1) the clashing of attack and defense as games of strategy and reveal that (2) attack benefits from mismatching its target's level of defense, whereas defense benefits from matching the attacker's competitiveness. This suggests that (3) attack recruits neuroendocrine pathways underlying behavioral activation and overconfidence, whereas defense invokes neural networks for behavioral inhibition, vigilant scanning, and hostile attributions; and that (4) people invest less in attack than defense, and attack often fails. Finally, we propose that (5) in intergroup conflict, out-group attack needs institutional arrangements that motivate and coordinate collective action, whereas in-group defense benefits from endogenously emerging in-group identification. We discuss how games of attack and defense may have shaped human capacities for prosociality and aggression, and how third parties can regulate such conflicts and reduce their waste.

## 1. Introduction

Social conflict has been part and parcel of human history and exerts a range of effects that easily exceed imagination. Conflict is associated with the rise and fall of nations and large-scale migration flows, and interferes with individual life trajectories. Conflict destroys welfare and lives, creates collective imprints and out-group resentments that transcend generations, and can cause famine and the spreading of infectious disease. Conversely, conflict drives technological innovation, inspires art, and creates and destroys hierarchies. Indeed, throughout history, conflicts have revised established structures and divides, introduced new views and practices, and changed the social order of individuals and their groups.

Conflict can be about many things such as ownership, territorial access, status and respect, and what is right and wrong (e.g., Blattman & Miguel 2010; Bornstein 2003; De Dreu 2010; Deutsch 1973; Gould 1999; Rapoport 1960; Schelling 1960). Sometimes, these conflicts are about something all parties want but that only some can have (Coombs & Avrunin 1988). Examples include politicians competing for the same senate seat, rivaling research laboratories claiming the patent ownership of a potentially lucrative technology, and superpowers seeking world hegemony. Alternatively, conflicts emerge because some parties want something that others try to prevent from happening (Durham 1976; Miller 2009; Pruitt & Rubin 1986). Examples include revisionist states seeking to capture their neighbor's territory, activist rebels fighting elitist powerholders, companies launching hostile take-over attempts, and terrorists attacking civilian and military targets.

Conflicts among nonhuman animals typically have such a structure of attack and defense (Boehm 2009; 2012; Dawkins & Krebs 1979; Sapolsky 2017; Wrangham 2018). Human conflicts may be no different. In fact, more than two-thirds of the 2,000 militarized interstate disputes that emerged since the Congress of Vienna in 1816 involved a revisionist state and a non-revisionist state (De Dreu et al. 2016a; Gochman & Maoz 1984; Wright 2014). Likewise, 68% of community disputes involved a challenger who desired a revision and a defender who protected the status quo (Ufkes et al. 2014). Finally, people often perceive the other side as the threatening aggressor who leaves them no option but to aggressively defend

themselves, a psychological bias often fueled by leader rhetoric (Chambers et al. 2006; Plous 1985; Staub 1996).

Although conflicts often have an attack-defense structure, theory and research have rarely made a clear distinction between attack and defense (Lopez 2017; Pruitt & Rubin 1986; Rusch 2014a; 2014b; Wrangham 2018). We currently lack theory and research about the ways in which clashes between attackers and defenders evolve, about the neurocognitive mechanisms that scaffold attack and defense, and about the cultural institutions that groups use to attack their neighbors or to defend against enemy attacks. Accordingly, our aim here is to provide a framework of the structure, psychological adaptations, and institutional consequences of attacker-defender conflicts.

We proceed as follows. Section 2 presents a formal model of attack-defense conflict and identifies unique properties of games of attack and defense that are neither present in nor captured by canonical games of conflict that have dominated classic and contemporary conflict analysis and research. Section 3 maps these properties onto behavioral decision-making and delineates its underlying neurocognitive mechanisms. We show that psychological "biases" often viewed as a result of imperfect human cognitive architectures may be functional for either attack or defense. Section 4 considers games of attack and defense between groups of people. We argue that groups face greater difficulty motivating and coordinating a collective attack of out-groups than defending against outside enemies, and that an out-group attack requires distinctly different sociocultural mechanisms and institutional arrangements than in-group defense. Section 5 shows how games of attack and defense, both within and between groups, may have shaped human capacities for cooperation and aggression.

## 2. The structure of conflict

*Mankind spends great amounts of energy on injuring others, and on protecting against injury.*

— John Stuart Mill (1848, p. 147)

The study of conflict incorporates a broad variety of conceptual and methodological tools, is strongly interdisciplinary, and encompasses multiple levels of analysis between individuals, groups of people, and (coalitions of) nation states (e.g., Blattman & Miguel 2010; Bornstein 2003; Cunningham et al. 2009; Gould 1999; Humphreys & Weinstein 2006; Pietrazewski 2016; Rusch 2014a). At the same time, there is a growing consensus to view conflicts as situations in which individuals or groups cannot realize their preferred state when other individuals or groups realize their own preferred state (Coombs & Avrunin 1988; De Dreu 2010; Deutsch 1973; Kelley & Thibaut 1978; Pruitt 1998; Schelling 1960). This conceptualization of conflict as incompatibility of interests is adopted here as well.

At the core of this unifying approach is behavioral game theory, which offers stylized models of conflict (e.g., Camerer 2003; Hirshleifer 1988; Kagel & Roth 1995; Schelling 1960). Indeed, game theoretical models of conflict have been used extensively in the study of international tension and interstate warfare (Bacharach & Lawler 1981; Huth & Russett 1984; Jervis 1978; Snyder & Diesing 1977) to understand the group dynamics and cultural arrangements that create and fuel intergroup conflict (Abbink 2012; Bornstein 2003; Colman 2003; De Dreu et al. 2014; Lacomba et al. 2014), to investigate the neural networks and neuroendocrine pathways involved in cooperation and competition (Cikara & Van Bavel 2014; Decety & Cowell 2014; De Dreu 2012; Rilling & Sanfey 2011), and to model the evolution of human prosociality and aggression (Bowles & Gintis 2011; Garcia et al. 2014; Henrich & McElreath 2003; Nowak 2006; Rusch 2014a; West et al. 2007).

Almost invariably, these lines of inquiry rely on models of conflict in which opposing decision-makers compete for the same reason(s). Accordingly, these models are ill-suited to approach the attacker-defender conflicts between revisionist and non-revisionist states, rebels and elitist powerholders, terrorists and security officers, or progressives and traditionalists. In this section, we highlight the structural properties of attack-defense conflicts that set them apart from the canonical games of conflict that dominate classic and contemporary conflict analysis and research (for notable exceptions, see, e.g., Carter & Anderton 2001; Dresher 1962; Durham 1976; Grossman & Kim 1996; 2002). These structural properties have critical implications for the incentives to engage in conflict or not, for the predictions about conflict expenditures, and for the psychological and institutional mechanisms underlying aggression or appeasement.

### 2.1. Conflict as incompatible interests

In its most basic form, game theory models conflict as two decision-makers (or "players") each with two possible actions to choose from, with the action that maximizes one's personal gain prohibiting the counterpart from maximizing her own personal gain at the same time. Imagine, for example, two countries both seeking world hegemony and building up a nuclear arsenal to subordinate the other side and to protect themselves against the other side's possible aggression. Only one country can achieve world hegemony, and the other must lose (or both lose when the nuclear option is used). Or imagine two farmers trying to gain exclusive access to the same water source, or two scientists working on the same problem trying to publish their breakthrough first. Again, what is in each player's best interest is incompatible with what is in the other side's best interest, and each will achieve its preferred state – world hegemony, access to the water source, or claiming a scientific discovery – only when the other does not.

These and related situations of incompatible interests have been modeled with various games of strategy. In the classic

Carsten K. W. De Dreu is Professor of Psychology at Leiden University, affiliated with the Center for Research in Experimental Economics and Political Decision-Making (CREED) at the University of Amsterdam, and fellow of the Royal Netherlands Academy of Arts and Sciences. He is the author of more than 250 publications in the area of cooperation and conflict, including group judgment, creativity, and decision-making. For his research, he received several awards, including the Carol and Ed Diener Award in Social Psychology from the Society for Personality and Social Psychology, an Advanced Grant from the European Research Council, and the Spinoza Prize, the highest scientific award given by the Netherlands Organisation for Scientific Research.

Jörg Gross is Assistant Professor at Leiden University and affiliated with the Leiden Institute for Brain and Cognition. He received his Ph.D. from Maastricht University. His work focuses on the origins of human cooperation and conflict and the neuroscientific basis of decision-making. For his research on cooperation and conflict, he was awarded a Veni grant from the Netherlands Organisation for Scientific Research and received funding from the Gratama Foundation and the Leiden University Fund.

Prisoner's Dilemma, for example, two players can choose between two possible actions labeled "cooperation" (C) and "defection" (D), as shown in Figure 1A. Each player can obtain one of four possible outcomes, depending on the action choice by the other side and oneself. Mutual cooperation (CC) is more beneficial to both players than mutual defection (DD). However, mutual cooperation is unstable because each player can increase its gain by playing defection when the other player cooperates, and vice versa. Player 1 thus prefers DC over CC, whereas player 2 prefers CD over CC. Players cannot both realize their preferred outcome at the same time.

Defection in the Prisoner's Dilemma is psychologically tempting for two reasons: First, defection can maximize personal gain. Second, defection protects against exploitation attempts of the other party (Coombs 1973). These two reasons for choosing defection rather than cooperation are disentangled in two other well-known games of strategy, the Game of Chicken and the Stag Hunt or Assurance Game (see Fig. 1B; Camerer 2003; Kagel & Roth 1995). In the Game of Chicken, choosing D can maximize personal gain *if* the other side cooperates. Thus, choosing D does not protect against loss, as in the Prisoner's Dilemma, but can increase personal gain compared with mutual cooperation (CC). This captures the conflict between two farmers who desire to move their cattle into new territory that can feed only one farmer's herd. Conversely, in the Assurance Game, choosing D cannot maximize personal gain but can protect against the worst outcome that is obtained when choosing C, if the other party chooses D. This captures a situation in which players do not expect their counterpart to cooperate (viz., distrust) and act preemptively to protect against exploitation (Abbink & de Haan 2014; Böhm et al. 2016; Halevy 2016; Simunovic et al. 2013).[1]

Crucially, these three basic games have in common that they are symmetric. Switching positions (player A becomes B, and vice versa) should not change their choice of strategy, nor the reasons for choosing that strategy. As such, opposing players have exactly the same structural motives for choosing C or D – gamble for maximizing personal gain, or avoid any risk and protect against loss and exploitation (Messick & Thorngate 1967; Pruitt 1967; Pruitt & Kimmel 1977). As such, symmetric games fail to capture the conflicts in which some players choose defection to maximize personal reward and others choose defection to prevent loss and exploitation. In such conflicts, switching positions (A becomes B, and vice versa) does not necessarily change their action but certainly the reason for preferring a certain action. For example, a rogue state contesting a superpower's world hegemony competes to increase its territory and its position in the global world order, whereas the superpower competes to protect its territory and to defend its top-ranking status position. When the leaders of both countries change positions, they may still decide to compete but now for diametrically opposite reasons.

## 2.2. Modeling the game of attack and defense

The distinct psychological reasons to compete – maximize reward versus protect against exploitation – separate players into attackers and defenders. In such conflicts, attackers have a preference-ordering similar to that of the Game of Chicken, whereas the defenders' preference-ordering is that of the Assurance Game. An ordinal variant of this game of attack and defense, which we call the Attacker–Defender Game (AD-G), is shown in Figure 1C.

Two features of the AD-G set it apart from symmetric games of conflict. First and foremost, in the AD-G, CC is more attractive

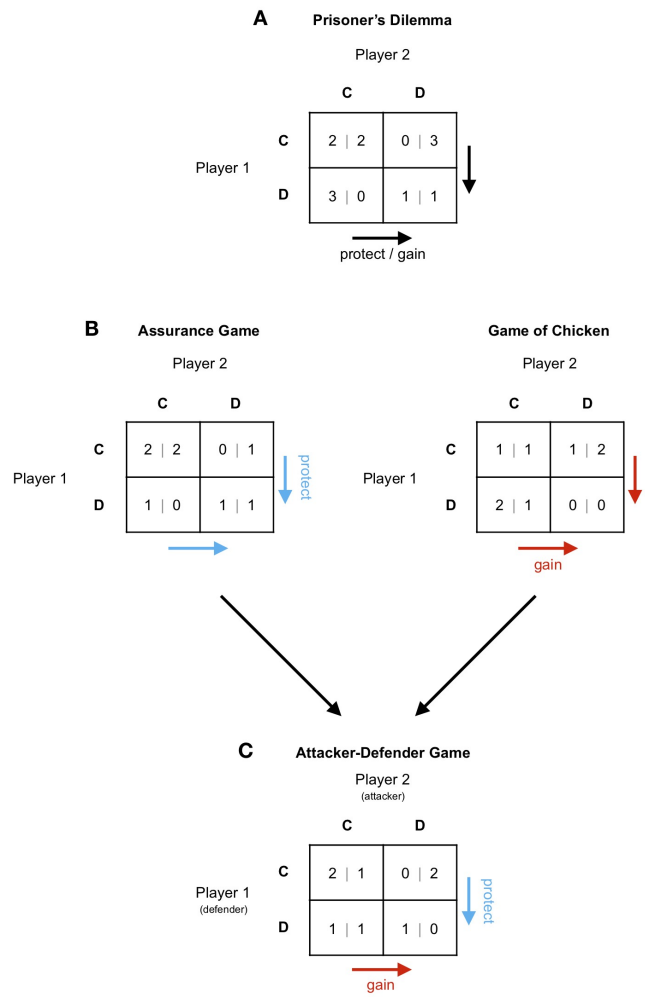


**Figure 1.** Games of conflict. (A) In Prisoner's Dilemma, both parties would mutually benefit from playing CC as opposed to DD. Playing D, however, can yield the highest *gain*. Further, playing C is risky, as it does not *protect* against exploitation. (B) In the Assurance Game, D *protects* against obtaining the worst outcome and guarantees a certain payoff. This situation reverses in the Game of Chicken, in which playing D can yield the highest *gain*, but does not protect against the risk of obtaining the worst outcome. (C) Combining the payoffs of player 1 in the Assurance Game with the payoffs of player 2 in the Game of Chicken leads to the Attacker-Defender Game. By playing D, player 1 can *protect* a loss with certainty (defend), while player 2 can gamble for a higher *gain* (attack).

to defenders than any other configuration of outcomes, whereas to attackers it is less attractive than the victory achieved when choosing attack (D) while the defender chooses to not defend (C). This fits the intuition that conflict may be triggered by relative deprivation and inequity aversion (see sect. 2.3). Thus, defenders benefit from peaceful interactions and compete to protect against exploitation, whereas attackers have an incentive to compete to maximize their personal reward. Relatedly, in the AD-G, collision (DD) is less costly to defenders than to attackers, which fits the defender's "home territory advantage"; whereas attackers need to overcome their rivals' defense, defenders only need to keep their attackers at arm's length (e.g., Galanter et al. 2017).[2]

Second, in AD-G, defenders prefer C > D when their attackers play C: Unilateral defense is costly. However, when defenders play C and thus are defenseless, attackers prefer D > C. Game-theoretically, the one-shot AD-G thus lacks a dominant strategy

for both the attacker and defender and has its Nash-equilibrium in mixed strategies.[3] This is also the case in some symmetric conflict games, such as the Game of Chicken, where defection maximizes reward when the counterpart cooperates and cooperation maximizes reward when the counterpart defects. However, in contrast to symmetric games with a mixed-strategy equilibrium, games of attack and defense have their equilibrium in an asymmetric matching-mismatching of strategies. Whereas it is in the defender's best interest to *match* its attacker's strategy (outcome DD or CC), it is in the attacker's best interest to *mismatch* its defender's strategy (outcome DC or CD) (e.g., Goeree et al. 2003). As an example, consider the Hide-and-Seek Game between a terrorist who seeks a target area where security officers will not look (mismatching strategy) and security officers surveilling areas where they think the terrorist is most likely to attack (a matching strategy) (Bar-Hillel 2015; Flood 1972; Steele et al. 2008; Von Neumann 1953; see also sect. 3.2 on social signaling).

Attacker-defender conflicts are often about an attacker's desire to improve on the status quo and a defender's desire to maintain and protect the status quo. The status quo defines a reference point (Kahneman et al. 1991; Samuelson & Zeckhauser 1988), with attackers trying to gain relative to the status quo and defenders trying to not lose relative to the status quo. To capture this, the AD-G can be transformed to a contest game (De Dreu et al. 2015), in which one player (henceforth, attacker) has to decide how much to invest in attack ($x$) out of a given endowment $e$ (with $0 \leq x \leq e$), while the other player (henceforth, defender) simultaneously decides how much to invest in defense ($y$) out of an equal endowment $e$ (with $0 \leq y \leq e$). If $x > y$, the attacker wins and obtains all of $e–y$. Added to the remaining endowment $e–x$, this leads to a total payoff for the attacker of $2e–x–y$, while the defender is left with 0. If $x \leq y$, the attacker appropriates nothing and the defender "survives," leading to a payoff of $e–x$ for the attacker and $e–y$ for the defender. This game is formally equivalent to a contest with a contest success function $f = x^m/(x^m + y^m)$, where $f$ is the probability that the attacker wins, with $m = \infty$ for $x \neq y$, and with the modification that $f = 0$ if $y = x$ (Dechenaux et al. 2015; Grossman & Kim 2002; Rusch & Gavrilets 2019; Tullock 1980).[4]

### 2.3. Summary and implications

Conflict theory and analysis mostly neglected models of attack and defense that emerge when states aggress their non-revisionist neighbors, when raiding parties attack adjacent communities, or when viruses battle with a host's immune system. Here we modeled such asymmetric conflicts as AD-G with a binary or continuous action space (see also Notes 2 and 4). The AD-G provides a stylized game-theoretic framework to formally analyze attacker's and defender's strategic choices, and to observe attack and defense in behavioral experiments.

Several potential extensions to our analysis are worth noting. First, our modeling of attacker-defender conflicts is limited to two-player conflicts and excluded multiplayer disputes in which more than two (groups of) individuals oppose each other. Multiplayer conflicts have an extra dimension of complexity because alliances among subsets of players can be forged that turn former foes into new friends and that change the power relations and payoff functions between rivaling factions. Second, as in any game of strategy, power is often asymmetrically distributed between antagonists (Bornstein & Weisel 2010; Choi et al. 2016; Durham et al. 1998; Hirshleifer 1991). Asymmetry in power can be modeled by inequality in resource endowments in the AD-G contest game and can dramatically change the motivation to attack or to defend. Finally, in dynamic settings, the position of attack and defense may change across time, depending on resources and conflict success, and repeated interactions can give rise to a *shadow of the future* that may increase the prevalence of conflict rather than promote peace (McBride & Skaperdas 2014; Skaperdas & Syropoulos 1996).

Our analysis, thus far, assumed strategic choices to be driven by the motivation to maximize gain (among attackers) and to avoid loss (among defenders). Humans are noteworthy for making social comparisons and strategic choices in conflict are also conditioned by the anticipated gain and loss *relative* to one's antagonist. For example, a wealth of research has shown how relative deprivation – having less than one's counterpart – drives strategic choices away from mutual cooperation and peace and toward conflict and competition (e.g., Halevy et al. 2010). Such social comparisons may differentially influence attackers and defenders (Chowdhury et al. 2018). For example, in the ordinal variant of the AD-G given in Figure 1C, both CC and DD would provide the attacker with less than the defender. Because of inequity aversion (Fehr & Schmidt 1999), attackers may be indifferent between CC and DD and prefer DC to CD. Conversely, attackers may anticipate guilt when their defection exploits a cooperating defender, and guilt aversion may inhibit the impulse to attack (Battigalli & Dufwenberg 2009; Dufwenberg et al. 2011; Ellingsen et al. 2010). To give one final example: Players sometimes value collective rather than personal outcomes (Bolton & Ockenfels 2000; De Dreu et al. 2000; Engelmann & Strobel 2004; Van Lange 1999), making cooperation rather than defection the preferred strategy among both attackers and defenders. In short, game theoretical analyses can provide a powerful tool to understand the structure of conflict, while behavioral experiments and psychological theory are needed to understand how people perceive, adapt, and react to these conflict structures. We explore this in the next two sections.

## 3. Psychological functions for attack and defense

*The rabbit runs faster than the fox, because the rabbit is running for his life while the fox is only running for his dinner.*

— Richard Dawkins & John R. Krebs (1979, p. 493)

Our model of attack-defense conflicts reveals structural properties that set them apart from symmetric games of conflict and that may have significant implications for conflict behavior and its underlying neurobiological and cognitive processes. In this section, we first review recent studies investigating the behavioral decisions attackers and defenders take. We then link these behavioral patterns for attack and defense to extant findings in neurobiological and psychological research regarding the neural networks, cognitive processes, and motivational biases related to cooperation and competition (Bazerman et al. 2000; Carnevale & Pruitt 1992; De Dreu & Carnevale 2003). In particular, we focus on neural networks involved in reward processing and threat detection (Molenberghs 2013; Rilling & Sanfey 2011), on literature linking aggressive hostility to overconfidence and biased perceptions of the rival's hostility (Ross & Ward 1995), and to feelings of superiority and tendencies to dehumanize opponents (Atran & Ginges 2012; Leyens et al. 2007).

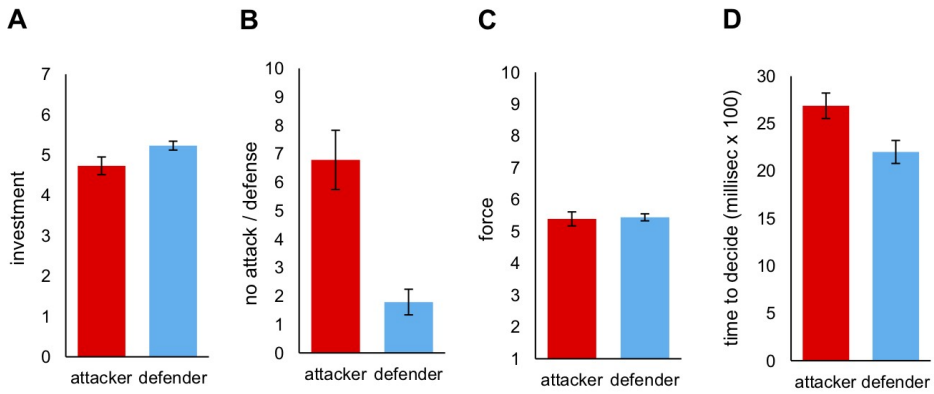


**Figure 2.** Behavioral strategies for individual-level attack and defense. Results from the aggregate of three incentivized experiments in which participants made 30–60 investment decisions in the role of attacker or defender, each time matched with a new partner. Shown are means ± SE, with $N = 85$ attackers and 85 defenders. (A) Overall investment (out of an endowment of $e = 10$). (B) Frequency of peaceful actions (no investment out of 30 trials). (C) Force of investments. (D) Time taken to decide.

### 3.1 Behavioral approach–avoidance in attack–defense conflicts

In general, people are loss averse; losses are more painful than commensurate gains are pleasurable (Kahneman et al. 1991; Kahneman & Tversky 1979; 1984), and people compete in prisoners' dilemmas more when their outcomes are framed as losses rather than gains (Andreoni 1995; Brewer & Kramer 1986; De Dreu & McCusker 1997; McCusker & Carnevale 1995; Sonnemans et al. 1998). Likewise, negotiators demand more and concede less when they focus on what they lose relative to their level of aspiration, rather than on what they gain relative to some rock-bottom resistance point (Bottom & Studt 1993; De Dreu et al. 1994; Kuhberger 1998).

The principle of loss aversion implies that attackers should compete less intensely than defenders (Chowdhury et al. 2018). Indeed, negotiation studies show that individuals who challenge the status quo (viz., attackers) engage in less domineering behavior, use punitive tactics less frequently, and are less successful than their counterpart who aims to maintain the status quo (viz., defenders) (De Dreu et al. 2008; Ford & Blegen 1992; Kteily et al. 2013). Furthermore, experiments using the AD-G contest game, outlined previously, show that attackers invested less than defenders (Fig. 2A; $F = 4.14$, $p = 0.044$) and less often decided to invest in attack than defenders decided to invest in defense (Fig. 2B; $F = 18.97$, $p = 0.001$; De Dreu & Giffin 2018; De Dreu et al. 2015; 2018). However, when investing, attackers and defenders used the same force; they invested about the same amount in attack and defense (Fig. 2C; $F < 1$).

Reward seeking has been linked to the neurobiological system of behavioral activation and approach, and the aversion of loss and punishment to the neurobiological system of behavioral inhibition and avoidance (Albert et al. 1993; Gray 1990). Conceptually, the behavioral activation system is triggered when the organism receives cues signaling rewards and controls actions that are regulating approach. Behavioral activation associates with positive emotions, such as excitement, hope, and optimism, in response to reward signals. It is modulated by the mesolimbic dopaminergic system and steroid hormones like testosterone (Ashby et al. 1999; Boot et al. 2017; Depue & Collins 1999; Eisenegger et al. 2011; Harmon-Jones & Sigelman 2001; Sapolsky 2005; 2017). Conversely, the behavioral inhibition system is triggered in response to anxiety-relevant cues and controls actions aimed at avoiding such negative and unpleasant events (Carver & White 1994; Elliot & Church 1997; Gray 1990). It associates with negative emotions such as fear, disgust and resentment, and, in the case of survival, relief.[5] Behavioral inhibition is modulated by the serotonergic pathway and stress-regulating hormones such as cortisol (Montoya et al. 2012; Nelson & Trainor 2007; Roskes et al. 2014; Sapolsky et al. 2000).

A first hypothesis emerging from this neuropsychological work is that attack and defense recruit distinct biobehavioral systems. In theory, attack should be associated with the release of the steroid hormone testosterone, mediated by the mesolimbic dopaminergic system, and activation in neural circuitries involved in the processing of rewards, such as the ventral striatum and the nucleus caudate. Conversely, defense should be associated with the release of cortisol and the recruitment of neural circuitries involved in threat detection and risk-avoidance such as the amygdala, the hippocampus, and the insula.[6]

A second hypothesis emerging from this work is that, with all else equal, attackers are disproportionally less successful than defenders. People invest less in attack than in defense, and the motivation to increase reward is weaker than the drive to avoid loss and defeat (Kahneman & Tversky 1984; see also sect. 3.3). Indeed, a negotiated settlement typically favors defenders rather than attackers (De Dreu et al. 2008), and our experimental results on attacker-defender contests showed that, defenders survived 66.4% of the contests (averaged over 60 one-shot rounds). A similar pattern emerges from archival analyses. When we analyzed success rates in the almost 1,500 militarized disputes between revisionist and non-revisionist nation states documented in the Correlates of War project (Gochman & Maoz 1984; Jones et al. 1996; Wright 2014), we found that only 25% were settled in favor of the revisionist state. Likewise, hostile takeover attempts in industry have a success rate less than 40% (De Dreu et al. 2016a). In short, attackers compete less intensely than defenders and have difficulty winning the conflict.

### 3.2. (Mis)Matching, deception, and social signaling

As discussed in section 2, a key property of asymmetric attack-defense conflicts is what we referred to as *asymmetric matching-mismatching* of strategies; attackers benefit from mismatching their defenders' level of competitiveness, whereas defenders should match their attackers' competitiveness. Studies of marital conflict regarding household chores and child care have documented a so-called demand-withdrawal pattern in which one spouse demands change in the other, who matches the intensity of the demand for change with an equally or more intense tendency to avoid interaction and discussion of the conflict issues (Kluwer et al. 1997; Vogel & Karney 2002; see also Mikolic et al. 1997). Behavioral experiments provide further support for this asymmetric matching-mismatching property. Specifically, we examined asymmetric matching-mismatching of strategies in 35 dyadic interactions across 40 rounds of Attacker-Defender contests. (For details on methods and materials, see De Dreu et al. 2016b.) Consistent with the principle of loss aversion,

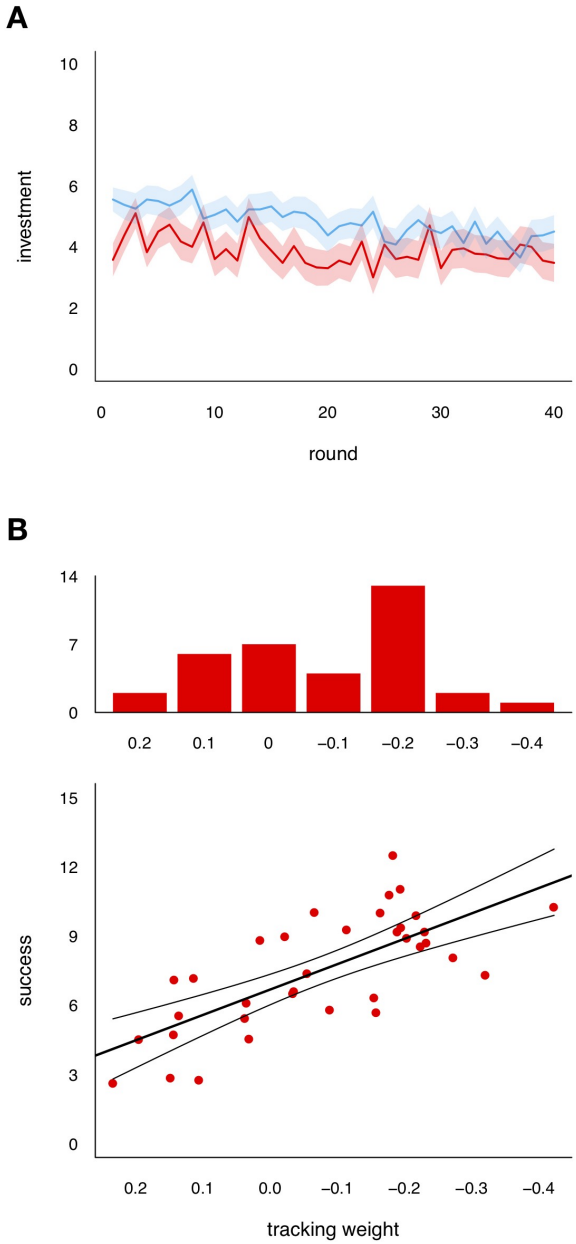


**Figure 3.** Strategic track-and-attack behavior. Results from an incentivized experiment in which participants made 40 investment decisions in the role of attacker and 40 investments as defender. In each 40-trial block, they were matched to the same partner and on each trial received full feedback. (A) Defenders (*blue*) invest more than attackers (*red*) on average (shown are means ± SE, with $N = 35$ attackers and 35 defenders). (B) Attackers are more successful when they condition attack on defenders' past behavior. The upper panel shows the distribution of regression weights for defenders' past investments (over the last three rounds) predicting attackers' expenditure on conflict. Negative values indicate larger attack expenditures when historic defense expenditure of the defender was low (and vice versa). Participants who systematically mismatch past defense expenditure are more successful (*lower panel*; final earnings on the *y*-axis; each dot represents one attacker).

we, again, found that defenders invested more than attackers (and both invested less at later rounds; Fig. 3A). We then calculated for attackers (defenders) the average investment of their opponent over the last three rounds and regressed investments in attack (defense) on this historical level of defense (attack). As shown in Figure 3B, attackers conditioned their behavior on their defenders' average investment: They attacked more when historical defense was low rather than high. And the more they engaged in such strategic forecasting, the more successful they were. Defenders, in contrast, invested more the higher their attackers' historical level of attack.

That attackers mismatch their defenders' strategy and defenders match their attackers' (expected) level of hostility have important implications for social signaling and deception among attackers and defenders. Precisely because both attackers and defenders lack a dominant strategy, they should be motivated to predict their rivals' future strategy and, at the same time, try to hide their own true intentions from their rivals. In games of attack and defense, defenders are motivated to signal strength and commitment to deter their rivals from attacking, whereas attackers are motivated to signal nonaggressiveness to lure their defenders into a state of (illusionary) safety (Slantchev 2010; Wheeler 2009). To paraphrase Dawkins and Krebs (1979): Whereas the rabbit runs faster and shows off its running strength with pride and confidence, the fox hides its true running capacity and instead feigns limpness.

### 3.3. Deliberate attack and spontaneous defense

Social signaling, attempts at deception, and accurate prediction of future events all require executive control and working memory. At the neuronal level, such a control network involves mainly prefrontal regions such as the inferior frontal gyrus, the dorsolateral and orbitofrontal gyrus, and the anterior cingulate (Aron et al. 2014; Braver 2012; Dosenbach et al. 2008; Posner & Rothbart 2007). A wealth of neuroimaging studies related these neural regions to risk assessment, to the inhibition of habitual responses and impulses, and to strategic planning and deliberation during decision-making (e.g., Aron et al. 2003; Coan & Allen 2003; Gross et al. 2018; Knoch et al. 2006a; 2006b; Mehta & Beer 2010; Peterson et al. 2008; Potegal 2012; Strang et al. 2015).

Although both attack and defense may be conditioned by executive control, there is reason to believe that attack recruits top-down control more so than defense. First, the attacker's task to mismatch its defender's strategy may require more controlled flexibility in thinking than the defender's task to reactively match its attacker's level of competitiveness. Indeed, in attack-defense contest games, attackers invested with greater variability (Fig. 2A) took more time to make their decisions (Fig. 2D; $F = 7.212$, $p = 0.008$), and they reported greater fatigue following the contest than defenders (De Dreu et al. 2018). Second, neuroimaging studies revealed greater activation in prefrontal control regions during attack than defense (De Dreu et al. 2015; Nelson & Trainor 2007; Siegel et al. 1999). Third, and, finally, there is evidence that reduced functionality of the control network affects attack more than defense. In one study, attackers and defenders performed cognitively taxing tasks prior to the contest. Results showed that, whereas defenders were not influenced by cognitive taxation, attackers made more aggressive investments when taxed rather than not (De Dreu et al. 2018). In another study, the functionality of the right inferior frontal gyrus (rIFG) was manipulated using Theta Burst Stimulation (De Dreu et al. 2016b). Again, defenders were not influenced and attackers more often invested when the rIFG was dysregulated, a pattern reminiscent of impulsive "high firing." However, when rIFG functionality was upregulated, attackers tracked their defenders' history of play more systematically and attacked when defenders were predicted to be weak rather than strong.

The hypothesis that attack is more controlled than defense can offer an explanation for the mixed findings on the role of

deliberation in public good provision games. Whereas some studies find that deliberation predicts more competition (e.g., Rand et al. 2012), others either find no relation between deliberation and competition, or find that deliberation predicts less competition (Bouwmeester et al. 2017). These studies invariably rely on symmetric games such as the *N*-person Prisoner's Dilemma, and the reason to compete can be to protect, to exploit, or a mixture of both motives. We propose that cognitive control and strategic deliberation play a stronger role when people compete for maximum reward than when they compete to avoid loss and exploitation (see also Simunovic et al. 2013). Games of attack and defense that disentangle motives of protection and exploitation may be a useful model to further understand when and whether deliberation promotes competition, or instead has little bearing on it.

### 3.4. Overconfidence and hostile attributions

People have a tendency to engage in motivated reasoning, searching and processing information that supports their goals and desires, while avoiding and downplaying information that is inconvenient or otherwise unsupportive (Jervis 1978; Kahneman & Tversky 1995; Ross & Ward 1995). Conflict theory and research have identified two types of motivated reasoning that are particularly problematic for conflict resolution and dispute settlement: overconfidence and hostile attributions. Overconfidence refers to overestimating one's relative strength (Deutsch 1973; Kahneman & Tversky 1995). Hostile attribution bias refers to overestimating malicious intent in others (Kramer 1995; Pruitt & Rubin 1986). In game-theoretic terms, overconfidence may be operationalized as an overestimation of the probability that one's rival plays C rather than D, and hostile attribution as an underestimation of the probability that one's rival plays C rather than D.

Overconfidence plays an important role in conflict spiral theory (Bacharach & Lawler 1981; Deutsch 1973). It argues that conflict escalates when and because (groups of) individuals believe they can win and emerge as the victor, for example, because they perceive themselves as relatively powerful. Indeed, overconfidence has been identified as a psychological precursor to conflicts such as the First World War (WWI), the Vietnam War, and the war in Iraq (Johnson 2004; Johnson & Fowler 2011; Van Evera 2003). In experimental war games, people who are overconfident about their expectations of success are more likely to attack (Johnson 2006), and, in bargaining games, higher overconfidence is associated with more competition (Neale & Bazerman 1985; Ten Velden et al. 2011).

Overconfidence has been linked to neuronal processes that we identified as involved in attack, including positive affect (Ifcher & Zarghamee 2014; Koellinger & Treffers 2015), the release of testosterone (Johnson 2006), and activation in reward processing areas such as the bilateral striatum (Molenberghs et al. 2016). In that sense, overconfidence may be functional to attack; it enables people to compete under risk (de la Rosa 2011; Johnson & Fowler 2011; Li et al. 2016). Or, as noted by Kahneman and Tversky (1995, p. 49): "Confidence, short of complacency, is surely an asset once the contest begins. The hope of victory increases effort, commitment, and persistence in the face of difficulty or threat of failure, and thereby raises the chances of success."

Defenders are unlikely to be overconfident. When confronted with potential attackers, overestimating one's strength and underestimating the rival's aggressive inclinations can be devastating. Rather, and perhaps therefore, defenders may be suspicious about their rivals' attack intentions, and their vigilant scrutiny may return biased impressions. In general, people more heavily weigh events that have negative, rather than positive implications for them (Pratto & John 1991; Taylor 1991). Also, person perception is influenced more by negative, rather than positive information (Fiske 1980). Automatic vigilance for negativity, such as for signals of the opponent's strength and malicious intent, may be accentuated during defense and elicit a hostile attribution bias (Kramer 1995; Pruitt & Rubin 1986; Waytz et al. 2014).

Hostile attribution bias can motivate (groups of) people to launch preemptive strikes aimed at neutralizing perceived threat and deterring one's rival from initiating attacks (Abbink & De Haan 2014; Bacharach & Lawler 1981; Halevy 2016; Jervis 1978). Preemptive strikes can provoke retaliation and, as such, create in defenders a self-fulfilling prophecy of strikes and counterstrikes that are wasteful and mutually destructive (Halevy 2016; Simunovic et al. 2013; Stott & Reicher 1998). In times of peace, on the other hand, hostile attribution bias could create a sustained and prolonged distrust in rivals that motivates investment in defense without apparent threats. There is some evidence that hostile attribution bias is indeed more prominent in defenders than attackers. In ideological conflicts between traditionalists who defend the status quo and revisionists who pursue change, traditionalists were more prone to polarize the two sides' attitudes and to attribute more extreme convictions to revisionists (Back 2013; Keltner & Robinson 1997; Robinson & Keltner 1996).

### 3.5. Feeling superior

People tend to believe that they are smarter, more moral, and less mean than others in general and their opponent, in particular. For example, experiments with professional negotiators, governmental decision-makers, and organizational consultants show that people view themselves as more constructive and as less destructive than their opponents (De Dreu et al. 2001), and such feelings of superiority are associated with increased hostility and an enhanced likelihood of future conflict (Babcock & Loewenstein 1997; De Dreu et al. 1995; see also Atran & Ginges 2012; Böhm et al. 2018; Ross & Ward 1995).

Feeling superior may be especially functional for attack. Attack means that targets may be harmed, subordinated, and exploited. Imposing such negative externalities onto others is generally inhibited by empathy (Batson 1998; Decety & Cowell 2014; Lamm et al. 2011), guilt aversion (see sect. 2.3), and social norms such as the "do-no-harm principle" (Baron 1994; Mill 1848). In recent experiments, we found that attackers with stronger other-concern, indeed, invested less in attack than attackers with weaker other-concern. Investment in defense, on the other hand, was not conditioned by other-concern (De Dreu et al. 2018). Feeling superior may reduce other-concern and provide the psychological justification for attacking others; it lowers the bar for using violence as a "means to an end" (Rai et al. 2017).

### 3.6. Summary and implications

The available research on competition in games of attack and defense permits three conclusions. First, attackers are less likely to compete than defenders and attackers invest fewer resources in fighting than defenders. Second, and possibly because of

these behavioral asymmetries, attackers are disproportionately less successful than defenders. Across a range of settings, from laboratory experiments to private sector competition to interstate warfare, we observed an attacker success rate averaging around 30% (De Dreu et al. 2016a). Thus, the prevention of exploitation is more likely than subordination. Third, we reviewed evidence that, consistent with the game structure of asymmetric conflicts, attackers are likely to mismatch the level of competition in their defenders, whereas defenders reactively match their attackers' competitiveness.

We have linked these distinct behavioral patterns to a range of neurocognitive functions underlying human attack and defense. We suggested that attack elicits behavioral approach and is associated more with prefrontal networks in the human brain than defensive behavior. Defense elicits behavioral avoidance and vigilant scanning for threat, and may be more automatic and intuitive than actions aimed at exploitation or profit maximization. Further, attacking and defending may markedly differ in overconfidence and risk-tolerance, hostile attributions, and feelings of superiority. Indeed, whereas overconfidence and feeling superior enable attack and may make people victorious against others, hostile attribution bias sustains defense and survival. Thus, one set of psychological biases may be the best response to the other set of psychological biases.

People possess both the dormant potential of overconfidence or dehumanization, on the one hand, and the potential for being vigilant and making hostile attributions, on the other hand. These abilities allow them to be both attacker and defender per situational requirements. At the same time, differences in psychological traits may differentially prepare people for these distinct roles in conflict. Taken to the extreme, the psychological characteristics of attackers merge into a psychological profile of people with high reward sensitivity, who are calculative and able to control impulses, are willing to accept risk, and lack empathy. Although these characteristics are usually temporary and triggered by contextual variation, when chronically activated, these characteristics merge into a profile reminiscent of trait psychopathy. Indeed, individuals labeled as psychopaths are impervious to the distress of others, lack fear of negative consequences of risky or criminal behavior, and demonstrate insensitivity to punishment (Book & Quinsey 2004; Hare & Neumann 2008; Meloy 1995; Patrick 1994). Conversely, the psychological characteristics of defense merge into a profile of people with high punishment sensitivity and risk aversion, who act intuitively, and are prone to making hostile attributions. When chronically activated, this profile is reminiscent of trait paranoia, a mental condition characterized by thought processes that include beliefs of conspiracy concerning a perceived threat toward oneself (Green & Phillips 2004; Saalveld et al. 2018). During the late 1960s, former U.S. President Lyndon B. Johnson, for example, found himself engaged in an intense political struggle around the ongoing war in Vietnam, and often complained that he could not trust anybody. By the end of his administration, Johnson "had become convinced that he was engaged in a life-and-death struggle, in which not only his foreign policy, but also his presidency … were at stake" (Kramer 1995, p. 127).

In as much as psychopathy and paranoia are considered dysfunctional pathologies, psychological science long considered the neurocognitive operations that impede constructive conflict resolution as manifestations of self-centered motivation and imperfect cognitive architectures (Bazerman et al. 2000; De Dreu & Carnevale 2003; Kahneman & Klein 2009). An alternative perspective, put forward here, is to view biases and motivated reasoning as adaptations to recurrent problems that humans repeatedly faced in the past (Cosmides & Tooby 1987; Fiedler 2000; Gigerenzer & Brighton 2009; Haselton & Nettle 2006; Tooby & Cosmides 1990). Key examples include overconfidence and feeling superior as adaptations to recurrent opportunities for attack, and hostile attributions as adaptations to recurrent threats of attack.

## 4. Intergroup games of attack and defense

> *Men, I now know, … fight for one another. Any man in combat, who lacks comrades who will fight for him, or for whom he is willing to die is … truly damned.*
>
> — William Manchester (1980)

Our framework, thus far, considered individual actors. To some extent, individual-level tendencies may operate also when groups of people engage in games of attack and defense. Indeed, entire groups can be overconfident about their attack being successful (Janis 1972) and feel superior to their rivals (Atran & Ginges 2012; Haslam 2006; Leyens et al. 2007). Likewise, entire groups can be in a state of persistent vigilance, prone to hostile attribution biases, and collectively dehumanize their rivals (Boyer & Liénard 2006). At the same time, however, intergroup conflicts have specific properties that are not present when unitary actors, such as individuals, engage in games of attack and defense. Specifically, in intergroup conflict, individuals within opposing groups have some discretion to contribute or not to the collective attack of out-groups or to join the collective defense of the in-group against an out-group threat (Bornstein 2003; see also Humphreys & Weinstein 2008; Radford et al. 2016). Because individual contributions to the intergroup conflict are not a given, groups make use of institutions like cultural rituals and sanctioning systems to motivate their members to fight or compete. Here, we review evidence that the structural properties of intergroup games of attack and defense require attacker groups, more than defenders, to create institutional arrangements that motivate and coordinate individual contributions to intergroup conflict.

### 4.1. Games of attack and defense between groups

Behavioral game theory has modeled intergroup conflict as "team-level games" in which group behavior depends on individual level preferences for in-group cooperation and competition (Aaldering et al. 2018; Abbink et al. 2010; 2012; Böhm et al. 2018; Bornstein et al. 1997; 2002; 2005; De Dreu et al. 2014; Halevy et al. 2008; Rapoport & Bornstein 1987). To illustrate, consider the team-level game of attack and defense between two three-person groups shown in Table 1. The group-level payoff matrix is based on the preference ordering in the Intergroup Chicken Dilemma (to model the attacker group's interests) and the Intergroup Assurance Dilemma (to model the defender group's interests) (Bornstein & Gilula 2003; Bornstein et al. 1996; De Dreu et al. 2016a). Within groups, however, individuals are faced with a classic public goods provision problem. In this example, within each group, individuals have a binary decision to contribute or not contribute personal resources to the group's fighting capacity. The attacker group wins when it has more contributors than the defender group; otherwise, defenders "survive." Individuals pay a cost $c$ when engaging in attack or defense. Assuming that the spoils from victory are divided equally among all members of the attacker group, it is in each team's best interest to have its members contribute, and it is in each

**Table 1.** Team-level game of attack and defense

| | $m_A$ | | | |
|---|---|---|---|---|
| $m_D$ | 0 | 1 | 2 | 3 |
| 0 | **6** \| 3 | *0* \| ***8*** | *0* \| *7* | *0* \| *6* |
| 1 | 5 \| 3 | **5** \| 2 | *0* \| ***6*** | *0* \| *5* |
| 2 | 4 \| 3 | 4 \| 2 | **4** \| 1 | *0* \| ***4*** |
| 3 | 3 \| **3** | 3 \| 2 | 3 \| 1 | **3** \| 0 |

$m_A$ = Number of contributors in the attacker group; $m_D$ = Number of contributors in the defender group.
*Note.* Entries indicate aggregate group outcomes (defenders *left*, attackers *right*). Defenders and attackers start with 1 "utility" point. When defenders do not defend, they earn 2 points each (e.g., they spend their time farming). Attackers and defenders pay a cost of 1 when attacking and defending, respectively. When $m_A > m_D$, the attacker group appropriates the resources of the defender group. When $m_A \leq m_D$, defenders survive and keep their earnings, whereas attackers have to pay the cost of attack without receiving any spoils from conflict. The upper triangle (*shown in italics*) of the payoff matrix constitutes attack-success. Numbers in **boldface** mark the best responses for each choice of the other group.

individual's best interest to not contribute in the hopes that others will.

Hence, individuals within both attacker and defender groups face a dilemma between what is good for their group (to contribute) and what is good for themselves (to not contribute). At the same time, the motivation to not contribute may be stronger in attacker than defender groups. Regardless of whether an out-group attack fails or succeeds, non-contributors in attacker groups earn more than contributors. This is different in defender groups. When a defense is successful, non-contributors in the defender group earn more than contributors. But when in-group defense fails, all members lose regardless of whether or not they contributed. Thus, in defender groups, individual interests are by definition more aligned than in attacker groups because they share a common fate when they lose.

As in individual attack-defense conflicts, intergroup conflict often arises because attackers seek an improvement over their status quo that defenders seek to protect. Thus, such conflicts are captured as team-level variants of the Best-Shot/Weakest Link Game (Chowdhury et al. 2013; Chowdhury & Topolyan 2016a; Note 4) and the Intergroup Aggressor-Defender Contest (IAD-C; De Dreu et al. 2016a) in which individual contributions are modeled continuously rather than binary. For example, assume an equal number of members in the two rivaling groups, with $N = N_A = N_D$. Each member $i$ is endowed with $e$ from which she can contribute $g$ ($0 \leq g \leq e$) to their group's fighting capacity $C$ ($0 \leq C \leq Ne$). Individual contributions to the pool $C$ are wasted, but when $C_A > C_D$, the attacker group wins the remaining resources of the defenders, with the spoils divided equally among attacker group members and added to their remaining endowments ($e - g + (Ne - C_D)/N$). Defenders thus earn 0 when attackers win. However, when $C_A \leq C_D$, defenders survive and individuals on both sides keep their non-invested resources ($e - g$). Accordingly, the incentive to free-ride is stronger in attacker than defender groups because, in case of failure, attackers keep what they did not contribute, whereas defenders earn nothing regardless of their contribution.

### 4.2. In-group cohesion and social identification

The stronger alignment of individual interests in defender groups, along with its anchoring on avoiding defeat, has important behavioral and psychological ramifications. Public good provision experiments show that, when group interests are more salient than individual interests, people are more likely to cooperate (e.g., Brewer & Kramer 1986; De Dreu & McCusker 1997). And indeed, in Intergroup Attacker-Defender Contests, individuals free-ride less when defending than attacking (De Dreu et al. 2016a; Zhang et al. 2018).

Facing an outside threat also drives people together into close-knitted groups (Boyer et al. 2015; Hamilton 1971; Radford 2008; Willems & Van Schaik 2017) and increases group cohesion and social identification (Calo-Blanco et al. 2017; Gilead & Lieberman 2014; Schaub 2017; see also Correll & Park 2005; Roccas et al. 2008). Cultural tightness, a tendency for groups to adhere to and enforce strong norms, is associated with more frequent attacks from enemy groups (Gelfand et al. 2013), and in-group identification is stronger the more group and individual outcomes can be negatively influenced by out-groups (Bobo & Hutchins 1996; Quillian 1995; Weisel & Zultan 2016).

We tested the possibility that in-group identification is stronger among defender than attacker groups by probing in-group identification following a series of contest rounds in which individuals contributed to out-group attack or to in-group defense. As predicted, we saw stronger identification among defenders than attackers ($F[1, 24] = 14.71$, $p = 0.001$; Fig. 4A). Importantly, and fitting the idea that in-group identification is a response to threat, identification among defenders was a positive function of their rivals' average investment in attack ($r = 0.615$, $p = 0.001$; Fig. 4B) and strongly predicted investments in in-group defense ($r = 0.431$, $p = 0.035$; Fig. 4C). Identification in attacker groups was unrelated to their rivals' level of defense ($r = -0.136$, $p = 0.528$) and negatively related to investment in out-group attack ($r = -0.422$, $p = 0.040$; Fig. 4B,C). Perhaps the latter finding is related to the relatively low success rates (approx. 25%) that attacker groups achieve even when they invest a lot (i.e., people dis-identifying from unsuccessful groups).

If in-group identification among defenders is a strong function of the threat posed by its rivals, the absence of out-group enemies may undermine in-group identification and loyalty. Because this makes the group a potentially attractive target for exploitation, groups may benefit from a continuous reminder of outside danger, which Boyer and Liénard (2006) refer to as the priming of a "mental hazard-precaution system." Examples include leader rhetoric that elicits enemy images of threatening out-groups and alludes to the dangers associated with being unprepared and off-guard (Staub 1996; Sternberg 2003), or rites and rituals that allude to clues of possible danger (Watson-Jones & Legare 2016). Modern-day rituals, such as military drills and mock battles, may similarly activate the mental hazard-precaution system. In addition to practicing and showing off strength and commitment to enemy states, they signal to their constituent audiences that the world is a dangerous place with untrustworthy neighbors.

### 4.3. Motivating contributions to out-group attack and in-group defense

Within defender groups, the stronger interest alignment and concomitant in-group identification can help resolve the individual tension between doing good to oneself (to not contribute) and doing good for the sake of the group (to contribute) in favor of the latter. This has a twofold implication. First, in-group cooperation emerges more spontaneously in defender rather than

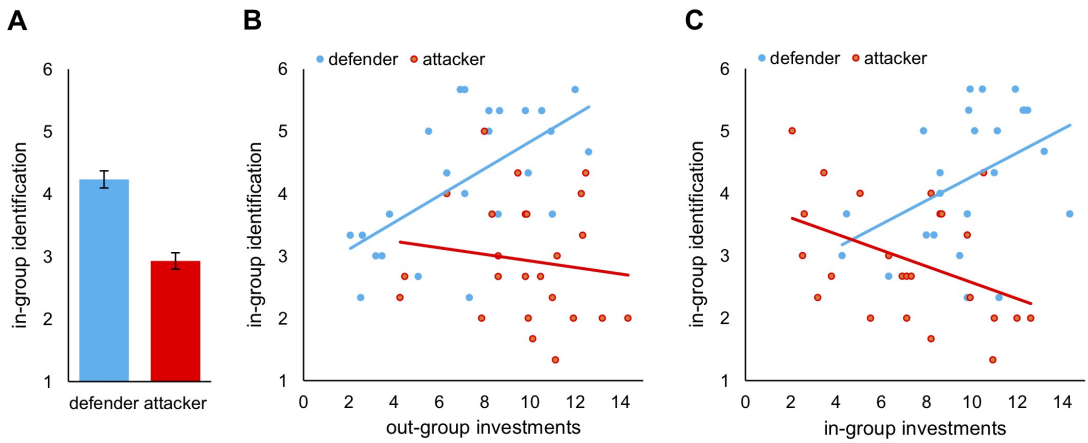


**Figure 4.** In-group identification. (A) In-group identification was measured after 10 investment rounds with one item: "I felt part of a group and identified with my colleagues" (1 = *not at all*, to 7 = *very strongly*). Ratings were averaged across members within defender and attacker groups ($N = 24$). (B, C) *Dots* show correlations (*blue* = defenders; *red* = attackers); *solid lines* represent best linear fit (*blue* = defenders; *red* = attackers). Data are based on unpublished results from De Dreu et al. (2016a, experiment 1).

attacker groups. Second, and therefore, attacker groups are in greater need of measures to motivate their members to contribute to group fighting and to deter them from free-riding on their team mates (see also Humphreys & Weinstein 2008).

Some measures that attacker groups use to motivate in-group cooperation directly aim to boost the otherwise fragile level of in-group identification. For example, attacker groups selectively invite friends to join a raid (Glowacki et al. 2016; see also Gould 1999; 2000), build strong bonds and friendships among its members (Macfarlan et al. 2014; Whitehouse et al. 2014), and engage in cultural rituals such as war dances that increase cohesion and commitment among its warriors (Fischer et al. 2013; Jackson et al. 2018; Lang et al. 2017; Whitehouse & Lanman 2014).

Other measures that attacker groups use are focused on deterring free-riding among its members. This includes the use of sanctions such as fines, physical punishment, gossip, and public shaming (Balliet & Van Lange 2013; Egas & Riedl 2008; Fehr & Gächter 2002; Henrich et al. 2006; Ule et al. 2009; Yamagishi 1986). An extreme example of deterrence aimed punishment is the execution of deserting soldiers who refused to actively participate in attacking the enemy, as was practiced during the trench warfare of WWI (Axelrod 1984). Experiments further show that punishment institutions indeed reduce free-riding and motivate individuals to contribute to their in-group's fighting capacity (Abbink et al. 2010, 2012; Bernard 2012; Gneezy & Fessler 2012). Consistent with the current hypothesis, we found that punishment institutions were used more in attacker rather than defender groups, and reduced free-riding (De Dreu et al. 2016a).

Sanctions do not have to be material or financial. For example, anthropologists have shown that belief in punitive gods, similar to peer punishment, not only promotes in-group cooperation, but also contributes to success in intergroup conflict (Atran & Ginges 2012; Johnson 2005; McKay et al. 2011; Norenzayan & Shariff 2008; Norenzayan et al. 2016; Purzycki et al. 2016). Resonating with the hypothesis that especially out-group attack groups benefit from religious beliefs is the finding that, across very different cultures, disadvantaged groups lacking religious spirit avoided aggression against their resource-rich and powerful counterparts, whereas disadvantaged groups with strong religiosity were less restrained and more likely to attack (Neuberg et al. 2014).

Although the literature on sanctions has largely focused on punishment, groups sometimes use moral, religious, or financial rewards to motivate members to contribute to out-group attack (Weinstein 2005). For example, to entice the English to join the Second Crusade in 1147 CE, Bernard of Clairvaux wrote: "Take up arms with joy and with zeal for your Christian name… O mighty soldiers, O men of war, you have a cause for which you can fight without danger to your souls; a cause in which to conquer is glorious and for which to die is gain. But to those of you who are merchants, men quick to seek a bargain, let me point out the advantages of this great opportunity … the reward is great" (Brundage 1962, pp. 92–93). An interesting avenue for future research is to examine the effectiveness of promising rewards relative to the threat of punishment in motivating people to contribute to out-group attack and reduce free-riding during out-group attack. Indeed, Doğan et al. (2018) varied the distributions of possible earnings from victory within groups and found that privileged group members pushed for higher aggression against out-groups than disadvantaged group members.

Apart from group cohesion and reinforcement, leaders and institutions sometimes try to motivate attack by transforming group members' beliefs about the conflict structure. Putting attackers into a defensive mindset not only switches the reference point from a potential gain to a looming loss, but also exploits that collective defense is a shared interest and perceived as morally superior, while attack faces a more severe free-rider problem and is more likely to be seen as morally devious. Indeed, to change attackers' belief of the "game they are playing," leaders and societies sometimes use propaganda – selective, misleading, and emotionally laden information aimed at creating an illusionary, yet threatening, scenario of loss and exploitation.

Recent history provides telling examples of such propaganda aimed at suggesting that attack is needed for defensive reasons (Fig. 5). Large-scale genocides and "ethnic cleansing" in Rwanda and former Yugoslavia served powerholders' desire for economic and territorial expansion and were justified by depicting "the enemy" as vicious threats to the nation's moral and cultural heritage and a danger to the nation's sovereignty (Mgbeoji 2006; Staub 1996; Sternberg 2003). Between 1933 and 1938, Hitler justified Germany's aims at expansion with appeals to collective security, equality, and self-determination (Goddard 2015). The anti-Jewish propaganda after WWI and during Nazi

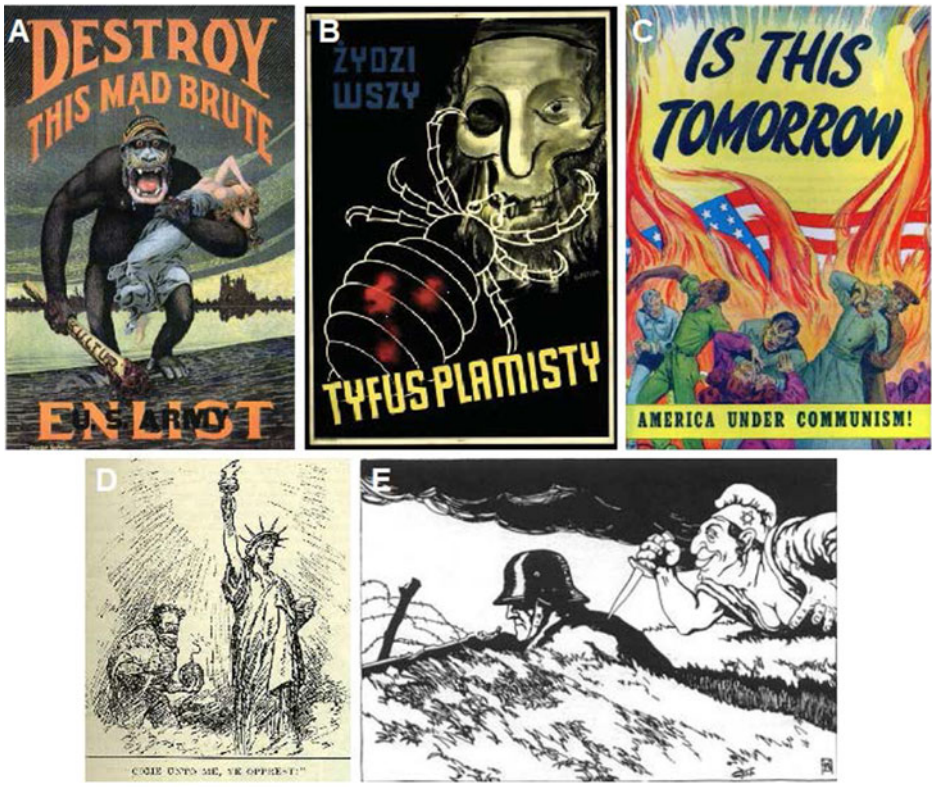


**Figure 5.** Examples of historical propaganda aimed at convincing the viewer of a threat to the status quo. (A) "Destroy this mad brute," a German soldier portrayed as a wild ape on the shore of America (WWI propaganda, ~1917). (B) "Jews, lice, and typhus," depiction of a deformed head behind the outline of a louse, trying to associate Jews with sickness and contagiousness (Nazi propaganda, Warshaw, ~1941/1942). (C) "Is this tomorrow," depiction of a burning American flag with fighting men in the foreground (anti-communist propaganda, 1947). (D) "Come unto me, ye opprest!" European Anarchist with dagger and bomb attempting to destroy the Statue of Liberty (American propaganda, 1919). (E) The depiction of a Jew backstabbing a German soldier at the front, illustrating the *Dolchstoßlegende*, a shared conspiracy theory during and after the Weimar Republic that the German defeat in WWI was caused by betrayal from inside (Austria, 1919).

Germany often portrayed Jews as scheming, backstabbing, and sneaky characters that betray and exploit "good patriots" when not on guard. A common theme of war propaganda is the creation of an imaginary attack on safety, health, and social order. During the Gleiwitz Incident in 1939, for example, Himmler's Schutzstaffel disguised as Polish attacked themselves. It allowed Hitler to frame his invasion of Poland the next morning as a defensive reaction to this attack.

### 4.4. Coordinating matching-mismatching of team attack and defense

By virtue of the common bad and the endogenously emerging in-group cohesion and identification, individuals within defender groups are aligned both in their self-sacrificial contributions to in-group defense and in their psychological orientation toward enemy threat. In general, such alignments enable tacit coordination on a shared focal point of not losing (Halevy & Chou 2014; Schelling 1960; Van Dijk et al. 2009). In contrast, coordination should be more difficult in attacker groups. Next to the pertinent problem of motivating members to contribute, groups should attack when their target's in-group defense is low rather than high (per the principle of matching-mismatching of strategies). Thus, attacker groups not only need to motivate its members to contribute the proper force, but also need to coordinate the timing of their attack. Our experimental results indeed show that attacker groups not only invested less than defenders across contest rounds (Fig. 6A), but also that the variance in investments across rounds was substantially higher in attacker than defender groups (Fig. 6B).

In the absence of endogenously emerging norms and focal points on which to coordinate, attacker groups need explicit coordination mechanisms to align individual actions. Examples of such mechanisms include pre-decision communication, deferring to a leader, and clear command hierarchies. Indeed, public good provision experiments have shown better coordination when group members were able to communicate rather than not (Abele et al. 2010; Alvard & Nolin 2002; Janssen et al. 2011; Oprea et al. 2014). Likewise, groups coordinate better when they have a leader (Glowacki & Von Rueden 2015; Gross et al. 2016; Levati et al. 2007; Potters et al. 2007; Van Dijk et al. 2003; Van Vugt & De Cremer 1999). Leaders can set the example for others to follow, and such leading-by-example coordinates contributions and reduces the variance in contributions within the group (Gächter et al. 2013; Hermalin 1998; Loerakker & Van Winden 2017).

There is some evidence for the hypothesis that especially attacker groups benefit from explicit coordination mechanisms. In two experiments using the Intergroup AD-Contest Game, we compared a baseline treatment in which group members decide simultaneously to a sequential decision-protocol in which one member made a first move to attack or not, followed by the second member, and so on. Defender groups were unaffected by this variation in decision procedure. Attacker groups, however, were significantly better at coordinating their contributions when they made attack decisions sequentially and were, therefore, more often victorious (De Dreu et al. 2016a; Zhang et al. 2018).

### 4.5. Summary and implications

Modeling intergroup conflict as team-level games of attack and defense revealed three insights. First, group-level defense creates a common fate for defenders that is absent in attackers and augments cohesiveness and in-group identification. Second, and relatedly, in-group defense elicits sacrifice and is more tacitly coordinated than out-group attack. Third, because out-group

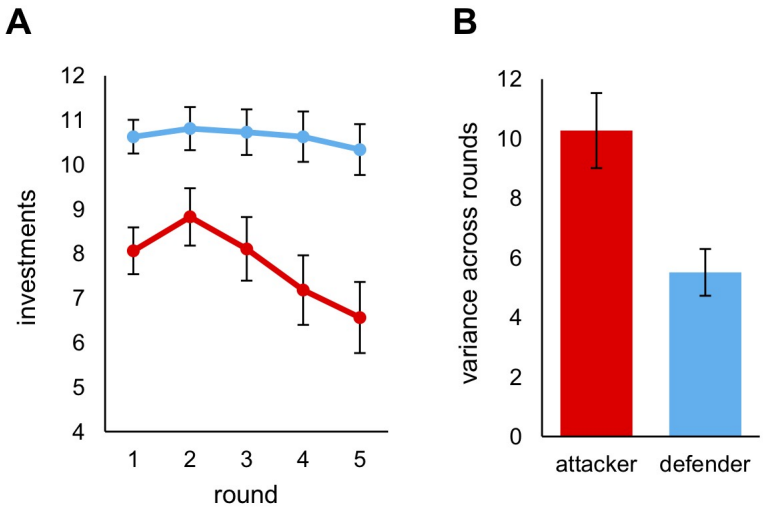


**Figure 6.** Group-level attack and defense across contest rounds. (A) Investments into defense and attack across the five contest rounds (displayed means ± SE). (B) Mean variance for investments into defense and attack across the five contest rounds (displayed means ± SE). Data are based on unpublished results from De Dreu et al. (2016a); baseline treatments of experiments 1 and 2 combined; $N = 46$ three-person attacker versus three-person defender groups.

attack is more vulnerable to both motivation and coordination failures than in-group defense, effective out-group attack requires sociocultural arrangements to deter free-riding, to motivate self-sacrifice, and to coordinate the timing of attack. Bonding rituals, sanctioning institutions, communication, belief manipulation, and leadership may be functional and emerge more for motivating and coordinating out-group attack than in-group defense. Possibly also, the greater threat of free-riding in attacker groups may promote steeper hierarchical power structures (Gross et al. 2016) compared with defender groups, and may have accompanied the transitions to the centralization of power, the emergence of "warrior class" specialization, and dedicated hierarchically organized military organizations when humans moved from hunter-gatherer societies to chiefdoms and states (Boehm 2009; 2012; Carneiro 1981; Earle 1987).

## 5. Conclusions and implications

*There is a constant struggle … between the instinct of the one to escape its enemy and of the other to secure its prey.*

— Charles Darwin (1873)

The present theory of attack and defense provides a threefold complement to existing conflict theory. We have argued that existing conflict theory is heavily focused on symmetric models. Symmetric conflicts, however, present a special class of conflict in which the motive to defend and the motive to attack and exploit are indistinguishable and present within each agent at the same time; the attacker is likewise defender, and vice versa. Asymmetric games of attack and defense, presented here, allow to tease apart these distinct motives. We have reviewed evidence that these new models are helpful in experimental studies to investigate the behavioral dynamics and underlying mechanisms involved when people seek victory and, alternatively, to protect against defeat and exploitation.

Teasing apart attack from defense can further reveal the functional relevance of a range of neurobiological, psychological, and sociocultural mechanisms for either attack or defense. As shown here, asymmetric conflict structures can help us understand when and why actors exhibit overconfidence, hostile attribution bias, dehumanization, and in-group identification. Our theory thus clarifies which of these psychological functions operate when and why.

By virtue of the conflict structure, the organization of attack and defense also requires distinctly different institutional arrangements to motivate costly contributions of group members in intergroup conflicts. Hierarchical power structures, punishment institutions, or belief manipulation and propaganda should play a more important role in motivating attack, while spontaneous coordination and voluntary acts of altruistic sacrifices should emerge more readily in the service of defense and protection (Rusch 2013).

Thus far, we have highlighted the structural and psychological similarities for attacker-defender conflicts across domains of conflict. Our analysis is assumed to apply to state-level actors in the political arena, to those involved in tribal warfare, to terrorist attacks and counter-surveillance, and to spouses who demand their partner to change. This is not to say that important differences exist across domains of conflict, with critical implications for our understanding of a particular conflict and our ability to predict its course. Indeed, although there is good neuroscientific evidence that human decisions are based on subjective value calculations (Bartra et al. 2013; Gross et al. 2014; Lebreton et al. 2009; Levy & Glimcher 2011; 2012), decisions are also influenced by frames, norms, heuristics, and psychological narratives (Gigerenzer & Brighton 2009; Gigerenzer & Goldstein 1996; Gigerenzer & Selten 2002; Kahneman & Tversky 1984; Kahneman et al. 1991; Simon 1956). For example, the do-no-harm principle discussed in section 3.5 may apply to physical harm much more than to attempts to change the status quo belief in society or to hostile takeovers in the industry (see also Fiske 1992; Fiske & Tetlock 1997; Heyman & Ariely 2004; Rai & Fiske 2011). Thus, although conflicts across domains and levels of analysis share key structural properties, they may not be psychologically homologous: Conflict domains differ in the rules and norms that define, to quite some extent, the subjective value and goals that decision-makers compute and pursue (Kelley & Thibaut 1978; Kelley et al. 2003). To understand a particular conflict and to predict its course, we not only have to take into account its structural properties, but also the social norms that apply, how conflict parties perceive "the game they are playing," and the concomitant subjective value that they derive from different actions.

Disentangling attack from defense opens up important avenues for new research, many of which we have discussed throughout. In section 2, we highlighted possibilities for extending the game-theoretical modeling of attacker-defender conflicts to

multiparty systems that afford shifting alliances and coalition formation, and repeated-play games that model the shadow of the future. In section 3, we highlighted the possible adaptive functionality of overconfidence and hostile attributions, and the potentially differential impact of social preferences and other-concern during attack rather than defense. And, finally, in section 4, we identified the possibility that opportunities for communication and leadership enable coordinated action that benefits out-group attacks more than in-group defense. Yet next to these specific research questions, our theory has more general implications for the understanding of the evolution of human social behavior and for conflict resolution. We discuss these in sections 5.1 and 5.2.

### 5.1. Games of attack and defense and the evolution of social behavior

Intergroup conflict has been an unfortunate constant throughout human history, from intergroup violence in ancestral hunter-gatherer societies to modern warfare (Campbell 1972; 1975). For group conflict, collective action and coordination are needed, and theoretical models suggest that cooperation and coordination may have indeed coevolved with conflict (Choi & Bowles 2007; Fu et al. 2012; Gross & De Dreu 2019b; Konrad & Morath 2012; Masuda 2012; Nowak et al. 2010; Rusch 2014a; 2014b; Traulsen & Nowak 2006). The coevolution of cooperation and conflict may explain why cooperation is often in-group bounded and parochial (Balliet et al. 2014; Gross & De Dreu 2019b; Bernhard et al. 2006; Bowles 2009; Choi & Bowles 2007; De Dreu et al. 2010; Efferson et al. 2008; Garcia & Van den Bergh 2011).

Much like contemporary conflict theory, models on the coevolution of conflict, coordination, cooperation, and institution formation assume symmetric games of conflict. For example, evolutionary agent-based simulations are mostly grounded in *N*-person Prisoner's Dilemma games or, in some cases, symmetric games of Chicken or Stag Hunt (Choi & Bowles 2007; Gross & De Dreu 2019b; Traulsen & Nowak 2006). If conflict is indeed the "midwife of altruism" (Bowles 2009), we should, however, take into account that conflict has often two faces, attack and defense, with distinct success functions and distinctly different social dynamics and institutional requirements. As noted, attacks on other groups require planning, coordination, and leadership and may have coevolved with complex communication skills, strategic thinking, and executive control. Group attack, more than defense, faces a free-rider problem as a result of imperfectly aligned incentives and benefits from establishing systems for sanctioning free-riders or steeper hierarchical social structures aimed at aligning incentives of group members to attack in coordination. Archeological and anthropological evidence suggests that coordinated out-group aggression increased with the transition from egalitarian hunter–gatherer bands with rather flat social hierarchy, to chiefdoms with steep hierarchies and specialized warrior classes (Boehm 2012; Earle 1987; Webster 1975). On the other hand, frequent threats from out-groups may have shaped the capacity for social support, solidarity, altruistic sacrifice, and heroism (De Dreu 2010; Rusch 2013). Defense also benefits from chronic vigilance, xenophobia, and spontaneous cooperation.

Hence, modeling conflict as an asymmetric game of attack and defense has important implications for how we should think about and model the coevolution of conflict and other human faculties, abilities, or biases. We suspect that defense may have been the midwife of altruism and xenophobia, while social institutions, hierarchical group structures, selective empathy, and dehumanization may have emerged in function of attack and exploitation.

Because the negative consequences of failed defense are stronger and more extreme than the consequences of failed attack, games of attack and defense create stronger selection pressures on defenders than on attackers (Brodie & Brodie 1999; Dawkins & Krebs 1979; Dugatkin & Godin 1992; Vermeij 1982). Throughout our review, we observed defense to be tougher and more strongly grounded in evolutionary older neural structures typically involved in fast and heuristic responding to threat. Even at the level of groups, in-group defense appeared spontaneously, tacitly well coordinated and modulated by endogenously emerging psychological functions such as in-group identification. In contrast, attack more often fails to be successful and may require evolutionary more recent brain structures involved in cognitive control and strategic deliberation. At the level of groups, out-group attack is successful especially when cultural rituals and institutions are invoked that combat free-riding and enable coordination of both the force and the timing of the attack. It is telling that even group-hunting predators engage social mechanisms to coordinate attacks. Wolves, once they have circled a moose, wait for the most senior wolf's move toward or away from the target, and then follow suit (Sand et al. 2006). In free-ranging African wild dog packs, the probability of rally success (i.e., group departure) is predicted by a minimum number of audible rapid nasal exhalations (sneezes), suggesting that some negotiation and voting shapes group-level decision-making (Walker et al. 2017). Our analysis here suggests that humans, as well, may need and develop cultural tools to expand and exploit.

### 5.2. Third-party intervention

In his *Principles of Political Economy*, John Stuart Mill (Mill, 1848/2008) argued that "it is the proper end of government to [take] measures as shall cause the energies now spent by mankind in injuring one another … to be turned to the legitimate employment of human faculties." He also identified "the aggressor [as] the person who first commenced violence by turning, or attempting to turn, another out of possession" (p. 7). Indeed, in any game of attack and defense, whether between individuals or groups, it is the attacker who initiates the conflict and the defender who reacts. Intervention aimed at preventing or de-escalating conflict typically aims to "transform the game" antagonists play (Halevy & Halali 2015; Nakashima et al. 2017) and, in the context of attacker-defender conflicts, could thus be targeted at strengthening the defender. Although this may de-motivate attackers, it may paradoxically tempt attackers even more because (unlikely) victory is now generating even more spoils (viz., Hirshleifer 1991). Alternatively, intervention may be aimed at reducing temptation in attackers. Indeed, when there are no thieves, there is no need to lock the door; when enemy soldiers disarm and return home, there is reduced need to mobilize one's army.

Our framework suggests two viable options that third parties have to reduce someone's temptation to attack. One option is to reduce the attacker's utility from winning, for example, by formal sanctions or condemnation. Many cultural, religious, and judicial practices have such functionality; religious leaders warn the greedy ones with images of fire and brimstone, communities imprison perpetrators, and politicians impose economic sanctions on rogue states. Experiments suggest that such sanctioning institutions can work well. When people anticipate or experience third-party punishment, they are less inclined to exploit others

and more likely to cooperate toward the establishment of common goods (Balliet & Van Lange 2013; Egas & Riedl 2008; Fehr & Gächter 2000; 2002; Gürerk et al. 2006). Thus, to reduce conflict, third parties could threaten to sanction attackers, and such threat should be effective when targeting the spoils of winning a fight, rather than targeting the mere attempt to attack others.

Another option available to third parties is to increase the utility that attackers derive from the status quo, for example, by donating money and food to relieve suffering, or to help people generate greater yield from the status quo by providing social and technological innovations (e.g., Carnevale 1986; Halevy & Halali 2015; Van de Vliert 1992; see also Fearon et al. 2009). This strategy is what John Maynard Keynes (1919) had in mind when he, as economic advisor to the British Government, argued against the Treaty of Versailles that settled WWI. In essence, Keynes's insight was that excessive war compensation payments demanded from post-WWI Germany would reduce the welfare of the status quo in Germany. Keynes (1919) worried that this would eventually push Germany to seek prosperity through (renewed) attack: "If we aim deliberately at the impoverishment of Central Europe, vengeance, I dare predict, will not limp."

Keynes' concern materialized with the Second World War (McDonough 1997) and resonates with studies in developmental economics and political geography showing that exogenous pressures and economic downturn correlate with the prevalence of conflict and warfare within and between societies (e.g., Allen et al. 2016; Buhaug & Rod 2006; Burke et al. 2015; De Juan 2015; Fjelde 2015; Prediger et al. 2014; Raleigh & Hegre 2009; Van de Vliert 2013). At present, the evidence on causality is inconclusive (Brunnschweiler & Bulte 2009) and limited to macro-level pressures and societal conflicts such as civil wars and interstate warfare. A possible hypothesis that emerges from our framework is that aid focused on improving the status quo can reduce the temptation to aggress and exploit others. An open question is which of the two options available to third parties – threatening to sanction the spoils of victory or providing aid to improve the status quo – is more effective in reducing attackers' temptation to aggress other (groups of) individuals.

### 5.3. Coda

Throughout human history, conflict restructured territories, alliances, and population dynamics, and may have shaped the biological and cultural capacities for motivated reasoning, for in-group cooperation and coordination, and for the deliberate exploitation of others. The twentieth century was characterized by a remarkable symmetric conflict between the Eastern and Western Bloc. That there was no large-scale fighting between the two superpowers has been attributed to the symmetric threat of nuclear annihilation, a stalemate driven by the looming possibility of mutual destruction. The post-Cold War era is once again predominantly characterized by asymmetric conflicts between terrorists and states, social uprisings and suppressing dictators, and information warfare in which foreign hackers try to destabilize enemy states. Asymmetric conflicts of attack and defense lack a pure-strategy equilibrium and are characterized by matching-mismatching dynamics and dishonest signaling on both sides. By the virtue of the structure of these conflicts, peace and stability should be harder to sustain in attacker-defender conflicts, resonating with the feeling that the world has become, once again, a more unpredictable place.

On the bright side, it has been argued that, at a global scale, conflicts become increasingly less frequent and that, when they occur, they are increasingly less violent (Pinker 2011; see also Falk & Hildebolt 2017; Mann 2018; Oka et al. 2017). If true, groups and societies no longer need to fear their neighbors and can invest resources in production and well-being rather than defense and protection. As shown here, doing so would make groups and societies not only prosperous but also vulnerable to greedy attacks by envious neighbors. As long as humans and their groups want both life and dinner, they are caught in games of attack and defense and both individuals and groups are bound to invest in injuring others and protecting against being injured.

**Acknowledgments.** The preparation of this article was supported by Leiden University, the Behavioral Economics Priority Grant from the University of Amsterdam, and Advanced Grant 785635 from the European Research Council to Carsten K.W. De Dreu. We thank Andrea Arciniegas, Lennart Reddmann, Andrea Farina, Michael Giffin, Nir Halevy, Zsombor Méder, Hannes Rusch, and Eric van Dijk for their comments and suggestions.

## Notes

1. Maximizing personal gain can be seen both in absolute terms and in terms of the relative advantage gained over the other side. Likewise, protecting against loss can be seen in both absolute terms and in terms of the relative loss vis-à-vis the other side (viz., inequity aversion). Accordingly, in empirical studies of cooperation and competition, maximizing personal reward has been referred to as *greed, spite,* or *appetitive competition.* Protecting against personal loss has been referred to as *fear, aversive competition,* and *betrayal aversion* (e.g., Ahn et al. 2001; Baumgartner et al. 2008; Bohnet et al. 2008; Coombs 1973; Messick & Thorngate 1967; Ten Velden et al. 2011). We refrain from using these terms because they may invoke associations with psychological states that may not fully and unequivocally explain the individual's choice to compete rather than cooperate.
2. A related game is the Inspection Game, which models situations in which a player D (e.g., a superpower) verifies the adherence of player A to some contractual obligation (e.g., arms reduction) that player A prefers to violate (Avenhaus et al. 1996; Nosenzo et al. 2013). As in the AD-G, attackers compete to maximize personal gain and defenders compete to protect against exploitation. And here, as well, attackers have an incentive to mismatch their defenders, who, in turn, have an incentive to match their attackers' level of competitiveness.
3. In the AD-G, with the example payoffs shown in Figure 1C, the mixed equilibrium strategy for the attacker is $EV(P2, C) = p + (1 - p) \times 1$; $EV(P2, D) = 2p + (1 - p) \times 0$; $1 = 2p$, hence $p = 0.5$; for the defender, it is $EV(P1, C) = 2q + (1 - q) \times 0$; $EV(P1, D) = 1q + (1 - q) \times 1$; $2q = 1$, hence $q = 0.5$.
4. A related contest is the Best-shot/Weakest-link game, which models situations where player A (e.g., a terrorist cell) can choose to attack on one or more battlefronts (e.g., subways and airports) that are defended by player D (e.g., security officers). Attackers win a bonus if they succeed in winning at least one battlefront; otherwise, defenders earn the bonus (Chowdhury & Topolayan 2016a; Clark & Konrad 2007).
5. A related theoretical framework is Regulatory Focus Theory (RFT; Higgins 1997; 2000), which examines the relationship between the motivation of a person and the way in which that person approaches the goal. RFT differentiates between a *promotion focus* on hopes and accomplishments, also known as *gains*, and a *prevention focus* based on safety and security, also known as *non-losses*. Thus, in current terms, a promotion focus may be stronger in attackers and a prevention focus may be stronger in defenders.
6. Behavioral inhibition does not necessarily lead to avoidance behavior such as fleeing or hiding. Studies in animal behavior, for example, have documented a range of proactive behaviors under threat such as predator inspection and mobbing (e.g., Dugatkin & Godin 1992; Griesser & Ekman 2005; Haberli et al. 2005). In humans, an example of such proactive responding to enemy threat is the use of preemptive strikes (see sect. 3.4). Such proactive responses in a defensive position may be particularly adaptive when fleeing is not an option.

# Open Peer Commentary

## The political complexity of attack and defense

Talbot M. Andrews, Leonie Huddy, Reuben Kline, H. Hannah Nam and Katherine Sawyer

Department of Political Science, Stony Brook University, Stony Brook, NY 11794-4392.
talbot.andrews@stonybrook.edu leonie.Huddy@stonybrook.edu
reuben.kline@stonybrook.edu hannah.nam@stonybrook.edu
katherine.sawyer@stonybrook.edu
https://you.stonybrook.edu/talbotmandrews/
https://you.stonybrook.edu/leonie/
https://sites.google.com/site/reubenckline/ https://www.hhannahnam.com/

doi:10.1017/S0140525X19000852, e117

**Abstract**

De Dreu and Gross's distinction between attack and defense is complicated in real-world conflicts because competing leaders construe their position as one of defense, and power imbalances place status quo challengers in a defensive position. Their account of defense as vigilant avoidance is incomplete because it avoids a reference to anger which transforms anxious avoidance into collective and unified action.

There is much to like in De Dreu and Gross (D&G)'s article, which develops a unique perspective on the differing psychological and neural processes of attack and defense. In a sweeping synthesis of findings from behavioral economics, neuroscience, cognitive science, and social psychology, the authors distinguish the unique properties of attack and defense. These differences generate several novel hypotheses and explain why it is easier to defend the status quo than bring about societal change.

Things get more complex, however, when delving further into two interrelated aspects of the theory: (1) the reality of real-world conflicts in which both sides construe their position as one of defense, not attack, and (2) the central role of anger in defensive intergroup conflict, which blurs the distinction between active attack and vigilant defense. The galvanizing and cohesive effects of anger help explain why political leaders justify attack as defense.

Consider first what may be the bigger problem for D&G, the tendency of leaders to "sell" attack as defense. The study of public opinion during war time provides clear evidence that Americans are most willing to engage in war when faced with a serious threat to national security, but, generally, reluctant absent immediate danger (Aldrich et al. 2006; Herrmann et al. 1999; Jentleson 1992; Jentleson & Britton 1998). National leaders may understand instinctively the psychology of attack and defense and reframe conflicts as defense even when they constitute attack. The Iraq War provides a compelling example. In the leadup to the war, the Bush administration advanced several specious or erroneous claims concerning Saddam Hussein's attempts to build nuclear weapons and connections to Al Qaeda that linked him to the 9/11 terror attacks (Althaus & Largio 2004; Feldman et al. 2015; Kull et al. 2003; Liberman & Skitka 2017). The administration persuaded a majority, if not all, Americans of Hussein's villainy and the necessity of a defensive attack on Iraq.

Collective action, which fits D&G's definition of attack as a challenge to the status quo, is also grounded in defensive psychology. In D&G's Attack-Defender Game, attackers and defenders have equal resources. But most political conflicts involve unequal power, complicating the subjective perception of attack and defense. For the less powerful, collective action can be viewed as a defense against existing ills such as sexual harassment (#MeToo), systemic racism (Black Lives Matter), or the erosion of living standards for the French working class (Yellow Vests). The status quo involves continued threat and loss for the less powerful consistent with evidence that institutions, norms, and leadership help structure the status quo and the value of its alternatives (Aldrich et al. 2006; Gelpi 1997). Furthermore, successful collective action requires strong group identities, a clear sense of efficacy, and a shared sense of grievance and deprivation, features described by D&G as emblematic of *defense* not attack (Klandermans & van Stekelenburg 2013; van Zomeren et al. 2008).

Our second concern is that D&G oversimplify the neuropsychology of attack and defense. They deserve credit for the development of a parsimonious model grounded in a basic neuropsychological dichotomy. In this scheme, attack is linked to reward, behavioral activation, approach, and positive emotions such as enthusiasm and hope, whereas defense is associated with punishment, behavioral inhibition, vigilance to threat, avoidance, and the negative emotions of fear and anxiety. This dichotomy is taken to its extreme when attack is equated with trait psychopathy and defense with trait paranoia. The binary distinction may aptly characterize reactions to potential gains and losses among individuals. Both animal studies (Crofoot & Gilby 2012; Kitchen & Beehner 2007; Rusch 2014b) and historical conflicts (Glowacki & Wrangham 2013; Rusch 2013), for example, Allied Europe versus the United States in World War II, suggest differences in the cost/benefit structure of intergroup attack versus defense. D&G wish to extrapolate to broader societal conflicts, however, that are inherently intergroup, changing defensive emotional and behavioral responses from anxiety to anger, a shift that is especially pronounced among strong group identifiers who believe the group can prevail (Mackie et al. 2000).

The link between anger and group defense poses an obvious challenge to D&G's model. Anger is a common response to group threat, especially among highly identified group members, and, in that sense, is integral to defense and collective action (Huddy et al. 2015; Van Zomeren et al. 2008). But problematically for D&G, anger is a defensive emotion that is linked to behavioral approach, occurs in regions of the brain associated with approach, and is linked to other positive (not negative) emotions (Berkowitz & Harmon-Jones 2004; Carver 2004; Harmon-Jones & Allen 1998; Harmon-Jones & Sigelman 2001). D&G acknowledge the prevalence of hostile attributions in defense, but they do not take the extra step to link hostility and anger. The absence of anger from their model is jarring and poses clear problems. They repeatedly characterize defense as avoidant and linked to behavioral inhibition, but anger does not fit that profile. It is an approach emotion that is associated with reward and behavioral activation, changing the essential psychological nature of defense. In that sense, D&G's characterization of defense as vigilant avoidance is overly simplistic.

The fundamentally different nature of individual and intergroup defense needs better explication in D&G's model. As they note, group cohesion intensifies under conditions of group defense, consistent with the well-documented effects of threat. This helps explain why group leaders transform attack into defense. The

creation of a threatening enemy is an effective way to foster group solidarity, generate anger, decrease risk aversion, and promote action. Leaders would not adopt this strategy if it led to risk aversion, avoidance, and inaction. As noted in several places in the manuscript, leaders specialize in the creation of threatening out-groups and enemies, which is difficult to explain without a more complex portrayal of the psychology of intergroup defense.

## Functional sex differences and signal forms have coevolved with conflict

D. Vaughn Becker[a] and Shelli L. Dubbs[b]

[a]Human Systems Engineering Program, Arizona State University, Mesa, AZ 85212 and [b]Department of Psychology, Benedictine University, Mesa, AZ 85201.
vaughn.becker@asu.edu shelli.l.dubbs@gmail.com
www.public.asu.edu/~loids/

doi:10.1017/S0140525X1900089X, e118

**Abstract**

Evolutionary theory makes further predictions about conflict. It predicts sex differences in the proclivity to attack and defend. It further suggests complementary biases in what we expect of the sexes. Finally, it suggests that the forms of human facial expressions of anger and happiness may have coevolved with the regularity of conflict as a means of signaling, bluffing, and defusing attack.

This is an exciting new chapter in game theory. We agree with the authors that we should "view biases and motivated reasoning as adaptations to recurrent problems that humans repeatedly faced in the past" (sect. 3.6, last para.) As evolutionary psychologists specializing in the social cognition of threat, we see at least three opportunities for testable hypotheses based on evolutionary theory: (1) sex differences in actions taken; (2) complementary biases in how the sexes are perceived; and (3) the ways in which facial expressions of prosociality and threat have coevolved with human perception to facilitate attack and defense.

Triver's (1972) theory of differential parental investment states that the lower investing sex (typically the male) will compete more intensely for access to the higher investing sex (typically the female). Consequently, a male's reproductive fitness (number of offspring produced in a lifetime) is often more variable than a female's. These selection pressures have endowed many species, including humans, with sexually dimorphic traits and behaviors. For example, men are taller and have greater upper body strength, are more competitive and aggressive (Archer 2004), and take greater risks compared with women (Wilson & Daly 1985). Such adaptations may help men directly defeat same-sex rivals, as well as acquire resources and status, which women find attractive in mates (Buss 1989). Taken together, men stand to benefit more from pursuing low-probability, high-reward strategies compared with women, because the potential gains in resources or social standing can increase a man's mating success. Women, on the other hand, tend to pursue a more cautious, self-protective strategy (Campbell 1999). For example, women report a greater fear of crime and victimization, and take actions to prevent or avoid victimization (May et al. 2010). On the basis of these differences, it can be predicted that male attackers will invest more in attacking (compared with females), whereas female defenders should invest more in defense (compared with males). These sex differences are easily tested and may already exist in previous work.

Although these hypotheses are straightforward, a more nuanced set of predictions arises with contextual moderators. Research reveals that men take more risks when exposed to attractive women (Baker & Maner 2008; Ronay & von Hippel 2010), and this relationship has been shown to be mediated by an increase in testosterone levels (Ronay & von Hippel 2010). In line with these results, men exposed to attractive female images should be more inclined to attack than normal.

Although men are generally the lower investing sex, fathers can, and often do, provide extensive care and resources to their offspring (Geary 2000). Thus, both genders have an incentive to compete for access to high-quality mates (Campbell 1999). Given this, when primed with a potential romantic rival, both men and women may increase their investment in attacking. Although, for women, this may be true only when primed with images of potential romantic rivals, as opposed to playing against a rival in person. While women behave negatively toward their "sexy peers" (Valliancourt & Sharma 2011), they tend to use indirect tactics, such as gossip, social exclusion, and derogation, as a means of aggression (Valliancourt 2013). If women are interacting with a physically attractive rival face-to-face, they might become more inhibited when playing the role of the attacker, because women may not be as comfortable with direct aggression. It is also possible that women may attribute more hostile intentions to a female rival, believing that she is more malicious than she actually is (a hostile attribution bias). Considering that women's competitive tactics tend to be subtle, this cognitive bias makes sense, because it might benefit women to err on the side of defending against potential attacks. Women who are pregnant or with young children may also be more prone to defense.

A related opportunity lies in a deeper consideration of how out-group stigmatizations show complementary sex differences. Men have historically been the attackers in intergroup conflict (van Vugt 2009; van Vugt et al. 2007), and it is the formidability of the males of the out-group that should calibrate our decisions to attack or defend. Some work suggests that activating a self-protective motive, plausibly the same as the BIS system the authors evoke as the defensive network, leads both men and women to devote more cognitive processing to out-group men than in-group men or women of either group (Becker et al. 2010). Arousing a self-protective motive also enhanced biases to see out-group members as enemies in a signal detection task, and a complementary bias to see in-group men as allies (even though 50% had detectable signs of threats; Becker et al. 2011). It thus seems plausible that priming self-protection will increase aggression against out-group men but decrease aggression toward in-group men, and perhaps enhance our tolerance for their acts of aggression against the men of the out-group. The methods of the target article could easily test such ideas.

Thus far, we have discussed new experimental directions for the conflict paradigm. At a theoretical level, the author's arguments are consilient with new ideas about the evolution of facial expressions that signal aggression (anger) and prosociality (happiness). Mounting evidence suggests that the angry expression masculinizes a face, whereas happiness has features of neotony and femininity; modern forms of these expressions may be partially scaffolded on preexisting gender recognition systems (Becker 2017). The conflict game could formalize the mechanism of such signal convergence. The form of the happy face also seems to have evolved to be highly

discriminable at short durations and great distances (Becker & Srinivasan 2014), whereas anger holds attention after it has been seen (Becker et al. 2014). Anger and happiness may thus be instrumental to the dynamics that play out within a conflict, suggesting additional research trajectories.

These possibilities are merely the beginning of what a more evolutionary approach can bring to the theses advanced here. Importantly, these points also emphasize that evolutionary psychology is not a set of "just so" stories, but that it makes testable predictions based on first principles of evolutionary biology.

# Reasons to strike first

William Buckner[a] and Luke Glowacki[b]

[a]Department of Anthropology, University of California, Davis, Davis, CA 95616; and [b]Department of Anthropology, Pennsylvania State University, University Park, PA 16802.
wvbuckner@ucdavis.edu glowacki@psu.edu
https://traditionsofconflict.com/about/ https://www.hsb-lab.org/

doi:10.1017/S0140525X19000840, e119

**Abstract**

De Dreu and Gross predict that attackers will have more difficulty winning conflicts than defenders. As their analysis is presumed to capture the dynamics of decentralized conflict, we consider how their framework compares with ethnographic evidence from small-scale societies, as well as chimpanzee patterns of intergroup conflict. In these contexts, attackers have significantly more success in conflict than predicted by De Dreu and Gross's model. We discuss the possible reasons for this disparity.

De Dreu and Gross (D&G) show that in games of attack and defense, it is in the defenders' best interest to match the attackers' strategy, whereas it is in the attackers' best interest to mismatch the defenders' strategy. They propose that coordination emerges more easily and spontaneously among defenders, who have stronger in-group identification and are expected to invest more resources in conflict. While their model usefully draws attention to the different payoff structures of attack and defense, it underestimates the advantages that attackers have in setting the stage of conflict and overestimates the ability of defenders to match the attackers' strategy. As a result, in contrast to predictions from their model, attackers usually fare better than defenders, having significantly less mortality and a greater rate of success. These patterns are found in both chimpanzees and in human groups. Unlike defenders who are pushed into conflict with only two strategies, conflict or flight, attackers choose whether or not to initiate conflict, as well as the time and location of their attack. This creates on-the-ground difficulty in matching strategies, because defenders may simply be incapable of this. The result is an extraordinarily high success rate for attackers compared with defenders.

Accounts from a vast number of human societies, as well as chimpanzees, show that attackers in intergroup conflict often seek to minimize risk to themselves and wait to attack until they have a strategic advantage (Gat 1999; LeBlanc 2016). When chimpanzees patrol the boundaries of their community, or during hunter-gatherer territorial skirmishes, individuals may flexibly switch between defense (or flight) and attack, depending on the balance of power between sides (Manson & Wrangham 1991; Wrangham & Glowacki 2012). Chimpanzees patrol the edges of their territory with hypervigilance, as D&G predict in the defense condition; however, when they encounter an enemy and have a significant numerical advantage, they can shift immediately to attack (Langergraber et al. 2017). The Ngogo chimpanzee community at Kibale National Park commit more intergroup killings per year than any other studied chimpanzee community, at more than twice the amount of the next closest group (Wilson et al. 2014). Yet they also have the lowest rate of mortality from intergroup killings of any recorded chimpanzee community (Falk & Hildebolt 2017). Ngogo chimpanzees patrol their territory in large groups, averaging over 17 individuals per party, and their attacking parties outnumber the victim's parties by 10 to 1 (Wilson et al. 2014; Wood et al. 2017).

The most common form of attack during intergroup conflict for many small-scale or uncentralized societies is the ambush, where attackers take advantage of the element of surprise, leaving defenders at a disadvantage in attempting an organized response (Gat 1999; Otterbein 2009; Wilson & Glowacki 2017). Walker and Bailey (2013) provide data on mortality rates for within and between-group violence across 44 traditional lowland South American societies. They note that out of 238 death events (such as duels, homicide, and raids), in only 5 (2%) did an attacker die (Walker & Bailey 2013). While D&G emphasize overconfidence as one psychological mechanism contributing to a greater willingness to attack, the ethnographic and primatological evidence indicates that attackers tend to act as though they have a keen awareness of their likelihood of success, among humans and chimpanzees alike.

Intergroup violence among the Andaman Island hunter-gatherers conformed to the general pattern found across many small-scale societies, where "The whole art of fighting was to come upon your enemies by surprise, kill one or two of them and then retreat" (Radcliffe-Brown 1922). Attacks would only occur when the attacking party was certain it could take the enemy by surprise (Radcliffe-Brown 1922). Attackers would retreat if they met resistance, while if the defenders were truly caught by surprise, they would often flee to save themselves (Radcliffe-Brown 1922). While D&G propose that coordination should emerge more readily among defenders, it is worth emphasizing how successful ambushes can lead individual defenders to defect quickly, fleeing for reasons of self-preservation. While there is theoretical and empirical work demonstrating the importance of in-group defense in promoting altruistic behavior (Böhm et al. 2016; Rusch 2013; 2014a), a greater consideration of the situations where flight is a higher payoff strategy for individuals than defense may also help elucidate the reasons why attacks have such high rates of success.

D&G emphasize that games of attack and defense create stronger selection pressures on defenders than attackers, because the costs of failed defense are often larger than the costs of a failed attack, but they underemphasize the benefits of successful attacks. There is evidence that successful warriors have increased reproductive success in some small-scale societies (Chagnon 1988; Glowacki & Wrangham 2015) and cultural incentives such as material benefits or increased social standing are commonly accorded to warriors across many traditional cultures (Glowacki & Wrangham 2013). Males can also benefit from alliances formed in a raiding context, cooperating with kin members or other close allies, and leveraging participation in raids to net better opportunities in marriage exchanges (Macfarlan et al. 2014; 2018). Chimpanzee males who patrol the boundaries of their communities and engage in coalitionary violence also tend to have

increased reproductive success (Gilby et al. 2013; Langergraber et al. 2017; Williams et al. 2004).

D&G make a unique contribution in examining conflict through asymmetric matching-mismatching strategies in games of attack and defense, and their framework provides possible avenues for further fruitful work. Investigating the conditions where defenders are incentivized to defect, and expanding on the advantages of the attack condition, may help reconcile the empirical record of high success of attack with D&G's prediction of easier coordination and greater success while defending.

# Emotions in attacker-defender conflicts

Patricia Cernadas Curotto[a,b,c], Eran Halperin[c], David Sander[a,b] and Olga Klimecki[b,d]

[a]Laboratory for the Study of Emotion Elicitation and Expression, Department of Psychology, University of Geneva, 1205 Geneva, Switzerland; [b]Swiss Center for Affective Sciences, University of Geneva, 1202 Genève, Switzerland; [c]Baruch Ivcher School of Psychology, Interdisciplinary Center (IDC) Herzliya, 46150, Israel and [d]Department of Medicine, University of Geneva, 1202 Genève, Switzerland.
patricia.cernadas@unige.ch eran.halperin@idc.ac.il
david.sander@unige.ch olga.klimecki@unige.ch
http://cms.unige.ch/fapse/EmotionLab/Members/patricia-cernadas-curotto/
https://www.eranhalperin.com/
http://cms.unige.ch/fapse/EmotionLab/Members/david-sander/
https://olgaklimecki.com/

doi:10.1017/S0140525X19000918, e120

**Abstract**

The distinction between attackers and defenders might help refine the understanding of the role of emotions in conflicts. Here, we briefly discuss differences between attackers and defenders in terms of appraisals, action tendencies, emotional preferences, and brain activities. Finally, we outline how attackers and defenders may differ in their response to emotion-based interventions that aim to promote conflict resolution.

In tense situations, emotions emerge that influence conflicts, shaping decision making and behavior (Bar-Tal et al. 2007; Halperin 2016). We propose that the distinction between attackers and defenders in asymmetric conflicts, as addressed by De Dreu and Gross (D&G), may inspire a new line of research that will help broaden the understanding of emotional processes and their implications in intergroup conflicts. More specifically, this commentary focuses on the impact that the role of being an attacker or a defender has on emotional experiences, appraisals, action tendencies, emotional preferences, brain activities, and responses to emotion-based interventions.

Appraisal theory offers a useful framework to shed light on why the perception of being an attacker or defender could have a differential impact on emotional experiences and action tendencies: It proposes that distinct combinations of cognitive appraisals (i.e., evaluations of an event) influence the emotions that are experienced (Sander et al. 2018; Scherer & Moors 2019). According to D&G, superiority and overconfidence are more typical for attackers, which would suggest appraisals of high certainty and high control. These appraisals, in turn, are usually related to feelings of anger, pride, and contempt (Fontaine et al. 2007; Lerner & Keltner 2000). For instance, anger predicts lower risk perception (Lerner & Keltner 2000; 2001), a bias that may facilitate competition in attackers. Conversely, defenders are described in the target article as vigilant. This could be associated with appraisals of low certainty and low control, which are typical of the emotion of fear (Lerner & Keltner 2000). Fear has been shown to elicit the perception that events are riskier (Lerner & Keltner 2000; 2001), which may explain the behavioral avoidance in defenders described in the target article.

In addition to distinct appraisals and action tendencies, attackers and defenders probably also differ in their emotional preferences (i.e., what people are motivated to feel). These differences might be explained by the instrumental approach to emotion (Tamir 2009; 2016), as well as by the motivation to feel emotions congruent with the self-image of being an attacker or defender. First, in line with the instrumental approach to emotion (Tamir 2009; 2016), groups prefer to experience particular emotions in order to attain their goals in contexts such as conflicts (Porat et al. 2016). In the target article, defenders and attackers are described as having distinct group-based goals: Whereas attackers aim to change the status quo, defenders aim at maintaining it. Thus, defenders may benefit from the motivation to feel fear because feeling threatened may reinforce in-group affiliation (Bar-Tal 2013), inciting them to invest more resources in collectively protective behaviors. In contrast, attackers may be more motivated to feel anger, which is associated with overconfidence and hostile action tendencies, which, in turn, facilitate fight behaviors. Moreover, overconfidence may be dysfunctional for defenders, because it may reduce their vigilance, thereby giving rise to devasting attacks.

Pertaining to brain functions, D&G argue that attack should recruit prefrontal top-down control more than defense does. This may seem contradictory to previous research showing that prefrontal brain structures and activities, which are also important for emotion regulation (Davidson et al. 2000), are related to less aggression and punishment (Giancola 1995; Klimecki et al. 2018; Nelson & Trainor 2007; Raine & Yang 2006). Factors that may matter in this context are the party's engagement in aggressive versus conciliatory behavior, as well as the intensity and temporal dynamic of a given conflict. It may thus be that defenders also show pronounced prefrontal brain activations when engaging in forgiveness behavior. In terms of intensity, usually more stressful situations are associated with reduced prefrontal top-down control in the brain and more activation in limbic structures (Arnsten 2009), which suggests reduced prefrontal activations in attackers and defenders during periods of intense and stressful conflicts. Whether conflict behavior and related brain activations in attackers and defenders can be influenced by interventions, and to what extent, remain to be tested.

Potential interventions that have been shown to provide beneficial effects in conflicts by changing emotions are reappraisal training (Halperin 2014; Halperin et al. 2013), indirect emotion regulation strategies (Halperin et al. 2011), and compassion training (Cernadas Curotto et al., in preparation). In light of the discussed differences between an attacker's and defender's emotions, certain emotion-based interventions might be more efficient, depending on whether the person identifies as an attacker or defender. Reappraisal is considered an emotion regulation strategy and consists of reinterpreting the situation, which triggers an emotion, to modulate its emotional impact (Gross 1998; 2001). For attackers, reappraisal training could therefore be used to reinforce the perception of the advantages of the status quo in order to

reduce their anger and their motivation to attack. In defenders, however, reappraisal training might be less efficient, as experiences of fear may interfere with the efficient use of reappraisal strategies. Besides reappraisal training, indirect emotion regulation can have beneficial effects for both attackers and defenders, because it can be tailored to target the appraisals that are constitutive of the most dominant emotions for each group (Halperin 2016). In defenders, this intervention may increase beliefs of defensive capabilities, thereby reducing their fear. In attackers, indirect emotion regulation may reduce contempt by altering the feeling of superiority. Another candidate for promoting conflict resolution could be compassion training, because it has the potential to overcome intergroup biases by cultivating benevolence toward all beings (Klimecki, in press). Compassion is defined as the feeling of concern for others' suffering, which is accompanied by the motivation to help (Goetz et al. 2010). Research from our team shows that compassion training can improve interpersonal relations in conflicts (Cernadas Curotto et al., in preparation). Because the target article described attackers with stronger "other concern" as investing less in attacks than that shown by attackers with lower other concern, strengthening compassion – which is conceptually close to other concern – may be a beneficial strategy for reducing attacks. Future research is needed to investigate these assumptions.

# The attack and defense mechanisms: Perspectives from behavioral economics and game theory

Subhasish M. Chowdhury

Department of Economics, University of Bath, Bath BA2 7AY, United Kingdom.
s.m.chowdhury@bath.ac.uk https://sites.google.com/site/subhasishmc/

doi:10.1017/S0140525X19000815, e121

**Abstract**

This commentary complements the article by De Dreu and Gross (2019) from the perspectives of behavioral economics and game theory. It aims to provide a bridge between psychology/neuroscience research and economics research in attack-and-defense by stipulating relevant literature, clarifying theoretical structures, and suggesting improvements in experimental designs and possible further investigations.

De Dreu and Gross (D&G) survey the literature on various behavioral aspects of attack-and-defense. While focusing mainly in the areas of psychology and neuroscience, they often use game-theoretic structure and results from behavioral economics to construct their review. The aim of this commentary is to complement their analyses with further relevant readings from behavioral economics and game theory, such that the baseline theoretical predictions may be clearer. We also provide some suggestions for the improvement and advancement of the experimental studies.

First, consider the game-theoretic aspects of attack and defense. Any conflict has the feature that the involved parties expend resources to win (or to avoid loss) and, irrespective of the outcome, the resources expended become sunk. Contest theory is the area of game theory that analyzes these situations. Two popular functions are used to determine the probability of winning of player $i$ in such games. Both take the form $p_i = \left(x_i^r / \sum_i x_i^r\right)$, where $p_i$ is the winning probability, $x_i$ is the resources spent by player $i$, and $r \geq 0$ is a parameter. When the outcome of a conflict is deterministic, that is, when a player who expends even slightly more resources than the rival wins for sure, then it is an all-pay auction (Baye et al. 1996), and $r = \infty$. However, when there is enough noise in the conflict and the outcome is not that straightforward, then it is modeled as a "lottery" (Tullock 1980), and $r = 1$. Individual attack-and-defense conflicts feature rank-order spillovers, and these two types provide very different results (see Baye et al. [2012] for all-pay auction and Chowdhury & Sheremeta [2011] for lottery). Although all-pay auction is covered at the end of section 2.2 in D&G, the important case of lottery is not considered.

Game-theoretic and behavioral models of attack and defense involve network externalities. But those for individual players and for groups are distinct (it is unclear in the article; see Note 4). See reviews of the theoretical and the experimental research by Kovenock and Roberson (2012) and by Dechenaux et al. (2015).

In individual multibattle contests, players fight on multiple battlefields with limited resources. There may be an attacker (terrorist) who tries to destroy such battlefields and a defender (government) who tries to save those. Here, the battlefields are connected as weakest link for the defender and as best shot for the attacker. Theoretical solutions for such models are provided in Clark and Konrad (2007), Arce et al. (2012), and Kovenock and Roberson (2018), among others. Experimental evidence (Kovenock et al. 2019) reveals that the theory better predicts subject behavior for the all-pay auction than for the lottery.

In a group conflict, group members exert resources that go through a production technology to create "group effort," which then determines the winning group. The technology can be additive (an easy version of this with all-pay auction *a la* Baik et al. [2001] is considered in D&G), or it can be something else. For attack and defense, the attackers face a best-shot technology, whereas the defenders face a weakest-link technology. Solutions are provided in Chowdhury and Topolyan (2016a) for all-pay auction and in Chowdhury and Topolyan (2016b) for the lottery; and the predictions are very different. Currently, there is no experimental research on such situations.

Note some important differences between psychology/neuroscience and economics experiments. Whereas economists stress use of a theoretical benchmark to understand behavioral mechanisms, others apply distinct techniques (e.g., functional magnetic resonance imaging). However, it is possible to combine both and arrive at a better understanding. The different disciplines independently arrive at the same results, which are not shared. For example, the high dispersion result (sect. 4.3 in D&G) is recognized as "overspreading" in economics (Chowdhury et al. 2014) but is never tested for an attack-and-defense frame.

For the individual attack-and-defense research, the article introduces a normal form of Attacker-Defender Game (AD-G). This (reproduced in Table 1) suggests that, although the theoretical equilibrium is in symmetric mixed strategies, in reality, defenders will choose "defend" more than attackers choose "attack."

Indeed, that turns out to be the case in the experiments, and it is explained through the idea of loss aversion. The subjects seemed to view the game as in Table 2, because loss-averse defenders weigh the negative amounts higher than their numerical values and play "defense" more.

**Table 1.** (Chowdhury) Proposed AD-G

| | Attacker | |
|---|---|---|
| Defender | Not attack | Attack |
| Not defense | 2, 1 | 0, 2 |
| Defense | 1, 1 | 1, 0 |

**Table 2.** (Chowdhury) Intended AD-G

| | Attacker | |
|---|---|---|
| Defender | Not attack | Attack |
| Not defense | 2, 1 | 2–2, 1–1 + 2 |
| Defense | 2–1, 1 | 2–1, 1–1 + 0 |

However, until the game is framed as in Table 2 explicitly, it is unlikely that subjects understand the game in such a way. They view the game as in Table 1, and choose the "defense" option because that is the riskless option.

Chowdhury et al. (2018) investigated the same but introduced a game with a lottery and continuous level of resources (instead of only two options). They essentially provided a general version of Table 2, controlled for risk, and still provided support for loss aversion resulting in higher expenditure in defense.

The simple AD-G in Table 1 cannot reflect the sunk cost of conflict, and that in Table 2 is an easy version of Chowdhury et al. (2018) with binomial space and an all-pay auction. Hence, it will be interesting to investigate a similar general setup as in Chowdhury et al. (2018) but with an all-pay auction – as implemented in the current study. It will also be useful to control for risk behavior along with other behavioral aspects, as detailed in the current study.

D&G successfully present the audience with an overview of attack-and-defense experiments (mainly) from psychology and neuroscience. We point out that to bridge those areas and behavioral economics/game theory, existing games that may be a better fit for investigating the same questions will allow theoretical benchmarks and behavioral mechanisms for predicted results. We also note that the current economics experiments do not exploit the techniques (e.g., response time, cognitive ability, eye tracking, neurological effects) used regularly in psychology and neuroscience. Using such techniques can provide broader knowledge. We hope for such a bridge to cross in the future.

# Collective action problems in offensive and defensive warfare

Agner Fog

Department of Information Technology. Technical University of Denmark, 2750 Ballerup, Denmark.
agfo@dtu.dk www.agner.org

doi:10.1017/S0140525X19000700, e122

**Abstract**

A collective action problem exists not only in offensive warfare, but also in defensive situations. The collective action problem is dealt with in the same way in offensive and defensive warfare: by strong leadership, discipline, rewards and punishments, strong group identification, strict religiosity, and intolerance of deviants. This behavior is explained in terms of evolutionary psychology.

De Dreu and Gross (D&G) investigate psychological, neural, and endocrinological differences between offensive and defensive behavior in two-person games. They are assuming that the same differences apply to many different domains of conflict, including warfare. However, conflicts between groups involve a collective action problem that is absent in conflicts between individuals. The article claims that the collective action problem is most relevant in offensive warfare while groups in a position of defense benefit from *endogenously emerging in-group identification*, *self-sacrifice*, and *tacit coordination*. But this in-group identification, self-sacrifice, and coordination require an explanation. A group defending its territory has a collective action problem, as has the attacking group, and the in-group identification is indeed part of a response to the collective action problem, as I will explain in the following.

A group trying to conquer territory from some other group has a collective action problem if the cost of fighting for the individual warrior exceeds his share of the group-level gain. Likewise, a group defending its territory against an invading enemy has a collective action problem if the cost of fighting for the individual warrior exceeds his share of the averted group-level loss. We cannot make any claim about which group has the highest collective action problem, unless we are able to calculate and quantify all fighting costs, gains, and losses.

In fact, the sharp distinction between offensive and defensive warfare is not always realistic in real world settings. Violent intergroup conflicts have often involved long sequences of escalation and retaliation where offensive and defensive tactics are used by both groups. Such ongoing conflicts may last for generations. The fighting groups may not remember how the conflict started, and they rarely agree on who was the initial aggressor.

A strong group cohesion is necessary for successful fighting. Various forms of group cohesion have been studied in many different research traditions under different names such as solidarity (Inglehart et al. 2006), collectivism (Minkov et al. 2017), cultural tightness (Gelfand et al. 2011), asabiyya (Turchin 2007), and regality (Fog 2017). There is general agreement that group cohesion is increased under conditions of fighting or collective danger, but few scholars have been able to explain the underlying psychological and cultural mechanisms. Recent research finds that the tendency to strengthen group cohesion in case of conflict or other collective danger is controlled by a psychological mechanism that may have evolved exactly because it deals with the collective action problem (Fog 2017).

Humans have a psychological tendency to become authoritarian and desire a strong leader in cases of collective danger. This psychological response mechanism has evolved, according to the so-called regality theory, because it helps overcome the collective action problem. The strong leader can reward brave warriors, punish cowards and defectors, impose discipline, and make strategic decisions. The members of such a group will develop a hierarchical political organization, strong group identification, strict religiosity, xenophobia, and intolerance of deviants. The cultural tendencies are quite opposite in groups that live under conditions of peace and security. People in the latter situation will not

support a strong leader who might be despotic and take advantage of everyone else. Instead, they will develop an egalitarian and tolerant culture. These two opposing cultural tendencies are called *regal* and *kungic*, respectively, for the warlike and peaceful conditions (Fog 2017).

The regal tendencies can be observed in groups in offensive and defensive situations alike because collective action is needed in both situations. The regal tendencies are readily arising when a group feels threatened by a militant neighbor group. Perhaps this is what D&G have observed as "endogenously emerging in-group identification." The regal tendencies may take a little longer time to cultivate in clearly offensive situations. A leader who detects an opportunity for territorial expansion may try to rally support for attacking a neighbor group. The territorial expansion is more advantageous for the leader than for his followers because, if successful, it enables him to rule over more people and a larger territory (Gavrilets & Fortunato 2014). There are many examples in history of leaders who have exaggerated or fabricated dangers to rally support for themselves and their imperial ambitions (Fog 2017). The history of empires shows that the territorial expansion process is self-amplifying. When the process is started, it may lead to a continuous growth of territory and a more and more regal culture, until the possibilities for expansion have been exhausted and the empire collapses under its own weight (Turchin 2007). The strong group identification and other regal traits are seen not only under defensive conditions, but also in the clearly offensive case of an expanding empire (Fog 2017).

# Attack versus defense: A strategic rationale for role differentiation in conflict

Catherine Hafer

Department of Politics, New York University, New York, NY 10012.
catherine.hafer@nyu.edu
http://as.nyu.edu/politics/directory.catherine-hafer.html

doi:10.1017/S0140525X19000888, e123

**Abstract**

Is there a strategic mechanism that explains role-contingent differences in conflict behavior? I sketch a theory in which differences in optimal behavior for attackers and defenders arise under initially symmetric conditions through the dynamic accumulation of differences in the distributions of traits in the subpopulations of potential opponents.

Distinguishing between defensive and offensive roles is key to understanding human behavior in conflict. De Dreu and Gross argue that differences in neurobiological responses between attackers and defenders indicate differences in motives and point to the need to model conflict as a game with asymmetric payoffs. This approach leaves a key question on the table: Why would humans evolve distinct neurobiological systems for attack and defense? A compelling explanation for role-contingent responses would derive differences in optimal behavior for attackers and defenders from symmetric initial conditions. Below I sketch a theory in which role-contingent behavioral differences arise from cumulative differences in the strategic environments of attackers and defenders that, themselves, are the result of equilibrium behavior in symmetric initial conditions.

To give a brief intuition: The process of gaining entry to a desirable group has predictable effects on the population characteristics of both those in and outside of the group. Thus, group membership is a noisy signal of relevant, unobservable individual traits (Spence 1973) and a strategically relevant difference between otherwise identical individuals. Conceiving of the possession of an asset as a kind of group membership, otherwise identical attackers and defenders will exhibit different behaviors in equilibrium, consistent with the empirical patterns summarized by De Dreu and Gross.

To see the underlying mechanism, begin by considering a classic, symmetric incomplete information war of attrition. Two individuals compete for a prize, such as a territory, by trying to outlast the other; when one of them quits, the other immediately claims the prize. Suppose for simplicity that both face identical costs of waiting but have potentially different valuations of the prize – whether because they possess different abilities to exploit it, or different alternative opportunities, or simply different preferences. Although each individual knows only her or his own value of the prize, both are drawn from the same population with a commonly known distribution of prize valuations, so they have identical suppositions about each other. Thus, they are in ex ante identical strategic situations, and in equilibrium, their optimal strategies will be identical, monotonically increasing functions of their own values of the prize (Bishop et al. 1978). It follows that the winner is always the contestant with the higher prize valuation.

Now suppose that this contest occurs in a larger environment in which many such contests, between many such (randomly chosen) pairs of contestants, are occurring simultaneously, and that after these contests are over, winners and losers can be readily distinguished by their possession of the prize. A loser may wish to try a second time to obtain a prize, against a different opponent. This second contest, however, is importantly different than the first one. The contestants are not in identical circumstances; the one who possesses the prize (the defender) is known to have won a previous contest, whereas the one who does not (the attacker) is known to have lost. Thus, each attacker knows that she or he faces an opponent who proved to be a higher type than some randomly selected individual faced previously; and each defender knows she or he faces an opponent who proved to be a lower type. Formally, the distribution of defenders first-order stochastically dominates the initial distribution, which first-order stochastically dominates the distribution of attackers.

This difference in the distributions of opponents affects the optimal strategic choices of contestants, relative to the initial fight, through two channels, both of which increase the optimal strategy for defenders and decrease it for attackers. First, suppose the defenders' strategies are the same as their optimal behavior in the first fight. The direct effect of the difference between the distribution of defenders (who have all won the first fight) and the initial distribution is that a randomly selected defender quits later than before. Thus, an attacker's expected cost of winning is higher than it was in the first round of fighting, and so her or his optimal strategy is lower. By the symmetric argument, the defender's optimal strategy is higher relative to the initial contest.

Second, consider the indirect effects via the changes in opponents' strategies just described. The decrease in the attacker's

optimal quitting time decreases the defender's expected cost of winning, which increases the defender's optimal quitting time. By the symmetric argument, the attacker's optimal quitting time decreases. Thus, the direct and indirect effects reinforce one another. After just one round of conflict, it is optimal in equilibrium for an individual to expend greater resources to defend possession of a prize than to obtain it, *holding constant her or his individual traits*.

The previous example supposes the conflict is a war of attrition, but the fundamental causal argument is the same regardless of the mode of conflict: When entry into the group must be won, membership in the group becomes a signal of traits promoting victory, and non-membership becomes a signal of traits promoting defeat. A trait or combination of traits that is not readily observed but that affects the individual's investment into winning (an example of the private-information "type" in the previous analysis) creates differences in the cost of competing or in the ability to exploit the prize.

This drives a wedge between the optimal strategies of fully rational and far-sighted defenders and attackers in every subsequent round of conflict, holding constant the individual's traits. Hafer (2005) shows that a closed system obtains a finite-time reachable steady-state in which strategies of defenders and (potential) attackers diverge so much that no conflict actually occurs. It is straightforward to extend the results to a case in which frequent shocks to the populations (such as those that would occur with birth and death) would result in some conflict in the steady-state, in which individuals with identical traits fight harder and are more likely to win when defending than when attacking. These results provide a rationale for the advantageousness of developing neurobiological responses not merely to conflict, but to developing role-contingent, distinct responses to being the attacker or defender.

# Resolving attacker-defender conflicts through intergroup negotiation

Nir Halevy

Graduate School of Business, Stanford University, Stanford, CA 94305.
nhalevy@stanford.edu
https://www.gsb.stanford.edu/faculty-research/faculty/nir-halevy

doi:10.1017/S0140525X19000694, e124

**Abstract**

The target article focuses on how attacker-defender conflicts are *fought*. This commentary complements it by considering how attacker-defender conflicts may be *resolved* at the bargaining table. I highlight multiple linkages between asymmetric intergroup conflict as modeled with the attacker-defender game and negotiation research and illustrate how the proposed model of attacker-defender conflicts can inspire new research on intergroup negotiation.

The target article makes important contributions to our understanding of asymmetric intergroup conflict by highlighting the physiological, psychological, social, and institutional processes that govern the behavior of parties who seek to challenge versus maintain the status quo. De Dreu and Gross (D&G) focus particularly on how "clashes between attackers and defenders *evolve*" (sect. 1, para. 4) and get decided. Here I complement their contributions by discussing how attackers and defenders may *resolve* their dispute at the bargaining table.

D&G allude to the negotiation literature in the target article. Intergroup negotiation constitutes a principal way through which groups settle their disputes (Aaldering et al. 2013; Halevy 2008; Putnam 1988; Steinel et al. 2009; Van Kleef et al. 2013). Intergroup negotiation is often used to resolve decades-long attacker-defender conflicts between revisionist forces and power-holders. Prominent examples include the 1990s negotiations in Northern Ireland and the Middle East, as well as the recent negotiations in Colombia, in which the Irish Republican Army (IRA), Palestinian Liberation Organization (PLO), and the Revolutionary Armed Forces of Columbia (FARC) represented the revisionist forces, respectively. Integrating D&G's game-theoretic approach with existing knowledge on negotiations can facilitate theorizing about and empirical investigations of intergroup negotiations in attacker-defender conflicts.

There are multiple links between asymmetric intergroup conflict as modeled with the attacker-defender game (AD-G) and research on negotiation. First, the structure of the AD-G emerges by modeling attackers' preferences using payoffs from the game of Chicken and defenders' preferences using payoffs from the game of Stag Hunt. Research suggests that individuals commonly use both of these classic games to mentally represent asymmetric intergroup conflict between revisionist and non-revisionist parties (e.g., the Israeli-Palestinian conflict: Halevy et al. 2006), as well as bilateral negotiations more generally (Halevy et al. 2012). Indeed, decision-makers sometimes mentally represent international conflict as an asymmetric game in which "we" (the in-group) play a game of Stag Hunt while "they" (the out-group) play a game of Chicken (i.e., the 1980s nuclear arms race: Plous 1985).

Second, interest misalignment is at the core of the AD-G and research on negotiations. Some of the motivating examples provided by D&G, such as "world hegemony" and "exclusive access" to resources (sect. 2.1, para. 1), suggest a zero-sum, step-level view of conflict outcomes, that is, a binary world in which only winning or losing matters and the margin of victory or defeat is irrelevant (Bornstein 1992). Consistent with the view that intergroup conflict is rarely zero-sum (Schelling 1980), negotiation research acknowledges that complex real-world negotiations entail continuous outcomes and encourages negotiators to renounce fixed-pie perceptions of conflict outcomes (De Dreu et al. 2000). Further, negotiation research suggests that differences in priorities across multiple issues present the parties in complex intergroup negotiations with opportunities to engage in mutually beneficial trade-offs to increase joint outcomes. Thus, for example, if one party prioritizes security arrangements and refugee resettlement over all other issues while the other party prioritizes territorial gains and access to natural resources (e.g., water) over all other issues, both parties can get more of what they want by making concessions on low-priority issues in exchange for concessions on high-priority issues.

A third opportunity for integration between the AD-G and negotiation research has to do with the propensity to initiate negotiation. Research on negotiation suggests that power asymmetry often inhibits disadvantaged parties' propensity to initiate negotiations (Bowles et al. 2007; Small et al. 2007). Research on intergroup negotiations in asymmetric conflicts, in particular, suggests that parties' willingness to come to the table depends on their beliefs

that they would be able to challenge versus uphold the status quo at the bargaining table. Specifically, members of disadvantaged groups show greater willingness to negotiate when they believe they would be able to tackle consequential issues earlier rather than later in the negotiation, whereas members of advantaged groups show the opposite preference for tackling consequential issues later rather than earlier in the negotiation (Kteily et al. 2013).

The proposed model of attack and defense conflicts facilitates the generation of novel hypotheses concerning parties' willingness to resolve their dispute through negotiations. For instance, D&G note that mutual cooperation "is more attractive to defenders than any other configuration of outcomes" and that mutual defection "is less costly to defenders than to attackers" (sect. 2.2, para. 2). Based on these features of the model, researchers can derive novel hypotheses such as these: (1) Defenders would be more likely to come to the table when negotiations are framed as enabling mutual cooperation. (2) Attackers would be more likely to come to the table when defenders signal unwavering commitment to expand all necessary costs associated with a head-on collision scenario.

Evidence from an asymmetric conflict between striking union members (the revisionist party) and government officials (the powerholding party) lends initial support to these predictions (Halevy et al. 2011). Consistent with the payoff structure of the AD-G, members of the revisionist party perceived their labor dispute as a game of Chicken, whereas members of the powerholding party perceived it as a game of Stag Hunt. The powerholding party signaled its willingness to pay the costs associated with a head-on collision by keeping all schools closed for more than 2 months as the teacher union went on an all-out strike. This intergroup conflict was eventually resolved via collective bargaining that was framed as an application of mutual collaboration.

Attackers and defenders often use intergroup negotiation to resolve their conflict. Integrating the insights and findings presented by D&G with insights and findings from negotiation research – including mental models and framing, signaling and communication processes, willingness to negotiate and willingness to engage in mutually beneficial trade-offs, and the relative influence of "hawks" and "doves" on negotiation processes and outcomes (Kahneman & Renshin 2007; Steinel et al. 2009) – has the potential to greatly enhance our understating of how attacker-defender conflicts are managed and resolved at the bargaining table.

# Symmetric conflicts also allow for the investigation of attack and defense

Joachim Hüffmeier and Jens Mazei

Institute of Psychology, TU Dortmund University, 44227 Dortmund, Germany.
joachim.hueffmeier@tu-dortmund.de jens.mazei@tu-dortmund.de
https://www.fk12.tu-dortmund.de/cms/psych/de/Home/Mitarbeiter/psychologie_ifp/Hueffmeier_Joachim.html
http://www.psych.tu-dortmund.de/cms/psych/de/Home/Mitarbeiter/psychologie_ifp/Mazei_Jens.html

doi:10.1017/S0140525X19000724, e125

**Abstract**

De Dreu and Gross argue that only *asymmetric* games allow the motives underlying defense and attack to be disentangled. However, the Prisoner's Dilemma Game Alt matrix (PDG-Alt matrix), a modified *symmetric* PDG, also allows these motives to be disentangled. Studies using the PDG-Alt matrix produced findings contradicting a central claim of De Dreu and Gross.

We applaud De Dreu and Gross (D&G) for their novel approach to modeling conflict and for their new theoretical framework on attacker-defender conflicts. Using an asymmetric matrix game as a vehicle to study players' motives to defend and to attack appears promising. However, we believe that a new theoretical framework should be able to integrate all pertinent prior research and, as such, should also account for known dynamics involved in symmetric conflicts. D&G (sect. 5, para. 1, emphasis added) argue that "symmetric conflicts … present a special class of conflict in which the motive to defend and the motive to attack and exploit are *indistinguishable*." Continuing their argumentation, De D&G (sect. 5, para. 1) stress that "asymmetric games of attack and defense … allow to tease apart these distinct motives." In contrast, we illustrate that the motives to defend and attack can also be disentangled in certain symmetric conflicts. Moreover, we show how broadening the focus to include symmetric conflicts enables the consideration of meaningful past insights – some of which contradict a central claim of the authors' framework.

We fully agree that asymmetric games allow the motives of defense and attack to be teased apart, but symmetric games can also provide an opportunity to do so. In fact, this has been successfully done in prior research (e.g., Insko et al. 1990; 1993; Schopler et al. 1993). Insko et al. (1990) developed the so-called PDG-Alt matrix, in which the cooperation and competition options are complemented with a third option, termed *withdrawal*. Choosing this withdrawal option protects from exploitation by the other party (i.e., when one player is willing to cooperate but expects competition from the other player) and it secures outcomes that are "intermediate to those obtained when both players cooperate or both players compete" (Insko et al. 1990, pp. 71–72; see Fig. 1). Importantly, providing a withdrawal option in the game allows researchers to tease apart the motives to defend and to attack.

To further illustrate how this is realized, and to explain why withdrawal is functionally equivalent to defense and why competition is functionally equivalent to attack, we draw on research on one-shot intergroup interactions with PDG and PDG-Alt matrices. Prior to the actual intergroup interaction, a substantial share of the groups expects competition by the other group, while a smaller, yet still substantial share of the groups expects cooperation (Insko et al. 1993; 2001). For groups expecting competition in one-shot interactions, withdrawal represents a defensive choice: By choosing withdrawal, they secure medium-sized outcomes, and they are unaffected by the other group's choice, but they cannot maximize their outcomes and they cannot achieve better outcomes than the other group. For groups expecting cooperation in one-shot interactions, competition represents a choice to attack: By choosing competition, they may be able to maximize their own outcomes, at the expense of the other group. Thus, for one-time interactions in the PDG-Alt matrix, withdrawal corresponds to defense and competition to attack. (Note, however, that the motives underlying withdrawal become less clear when parties engage in repeated trials with the PDG-Alt matrix. On the one hand, choosing withdrawal may still be a defensive choice. On the other hand, choosing

| Player B \ Player A | Cooperation | Withdrawal | Competition |
|---|---|---|---|
| Cooperation | 114 / 114 | 96 / 96 | 132 / 66 |
| Withdrawal | 96 / 96 | 96 / 96 | 96 / 96 |
| Competition | 66 / 132 | 96 / 96 | 78 / 78 |

**Figure 1.** (Hüffmeier & Mazei) Visualization of a PDG-Alt matrix (adapted from Insko et al., 1990, p. 71). As explained by Insko et al. (1990, p. 71), "the four corners of the matrix are a PDG matrix," whereas "the center column and row … represent safe (defensive or withdrawal) responses in which the choice of the other player is irrelevant."

withdrawal may be a tactical choice, such that players only resort to withdrawal following own competition in a prior trial to prevent costs from revenge.)

Considering that defense and attack can also be teased apart in symmetric conflicts has theoretical and empirical implications: The empirical evidence provided by Insko and colleagues using the PDG-Alt matrix contradicts the central claim by D&G that it is more difficult to motivate and coordinate collective action for out-group attack than for in-group defense. Research using the PDG-Alt matrix showed – for one-time interactions – that groups choose competition (attack) at least as often as, if not even more often than, withdrawal (defense; Insko et al. 1990; 1993; Schopler et al. 1993).

Other lines of research provide further evidence for this notion. For instance, studying groups' and individuals' offers in an Ultimatum Bargaining Game (UBG), Robert and Carnevale (1997) found that groups were particularly demanding. This demandingness cannot be explained by a defensive motive because demanding more than half of the available resources in a UBG does not represent a defensive choice.

Moreover, in negotiations, we also find a high prevalence of intergroup competition (e.g., Hüffmeier et al. 2018; Zerres & Hüffmeier 2011), which does not support the claim that group members need to be specifically motivated and/or coordinated for attack. In negotiations, the parties do not make discrete and simultaneous choices in one or several trials – like they typically do in matrix games – but, rather, discuss the course of action in iterative steps. Thus, there is a low risk of being attacked while being unprepared for it. Accordingly, there is also a low need to engage in defensive acts or preemptive strikes. Yet, even in negotiation contexts, a surprisingly high level of competition especially among groups remains prevalent (e.g., Hüffmeier et al. 2018; Zerres & Hüffmeier 2011).

In summary, both types of conflict – asymmetric as well as certain *symmetric* – allow the motives of defense and attack to be disentangled. Moreover, the predictions and findings for both conflict types differ from each other. Therefore, a framework on the mechanisms underlying defense and attack should be able to integrate both conflict types and reconcile these diverging findings. Going forward, we suggest considering the neglected line of research on symmetric conflicts and extending the scope of the proposed framework accordingly.

# Unraveling the role of oxytocin in the motivational structure of conflict

René Hurlemann and Nina Marsh

Division of Medical Psychology, University Hospital Bonn, 53105 Bonn, Germany.
renehurlemann@icloud.com n.marsh@gmx.de
http://renehurlemann.squarespace.com/welcome/

doi:10.1017/S0140525X19000785, e126

**Abstract**

Current psychological perspectives emphasize "attack" and "defense" as the behavioral mechanisms underlying conflict. Here, we extend this view by highlighting the relevance of pathological altruism and the neuroendocrine pathways associated with hostile behaviors. Specifically, we elucidate the modulatory role of the neuropeptide hormone oxytocin in motivating extraordinary levels of in-group commitment that can promote extreme behaviors and endure conflict with out-groups.

Across cultures, enduring conflict remains an integral part of human societies with often devastating consequences. In the twentieth century alone, almost 135 million individuals died from conflict, including 22 million murdered and 112 million killed in war (https://informationisbeautiful.net/visualizations/senseless-conflict-deaths-per-hour/). In their superb article, De Dreu and Gross focus on conflict by decomposing its underlying behaviors into "attack" and "defense." However, because humans are notable for the scope and variety of their prosocial behavior, which find no equal among any other species (Silk & House 2011), the nature of conflict may not be defined exclusively by antisocial behaviors. In this commentary, we therefore extend the view of De Dreu and Gross by highlighting the role of pathological altruism and its associated neuroendocrine substrates in the origin of conflict.

Altruistic and hostile behaviors are often viewed as diametrically opposed. However, given that both behaviors can take on different facets owing to variability in individual, motivational, and situational factors, attack and defense may not be uniformly linked to antisocial acts in the same way as altruism may not be universally prosocial (Marsh 2018). One such example is pathological altruism. In contrast to the positive outcomes associated with philanthropic altruism, in general, pathological altruism is thought to reflect extreme acts of selfless behavior resulting in negative consequences to the self and even to innocent others (e.g., family members), ranging from co-dependency to death (Oakley 2013). Prototypical examples include Christian martyrdom or burnout in health care professionals (Hurlemann &

Marsh 2016); even suicide bombings might be caused by pathological altruism when perpetrators of these acts believe they are behaving altruistically toward those who share their ideology (Habash 2011). Pathological altruism is deeply rooted in the evolution of cooperation and parochialism, characterized by in-group favoritism on the one hand and hostile and derogatory behaviors toward out-groups on the other (Choi & Bowles 2007). The formation and maintenance of in-group alliances are often enforced by internalized social norms (Bernhard et al. 2006; Gavrilets & Richerson 2017) and personally costly sanctions against defectors of these norms (Izuma & Adolphs 2013) – an inclination defined as altruistic punishment, possibly evolving to protect in-group alliances from erosion through selfish motives (Fehr & Gächter 2002). Thus, pathological forms of altruism may result from unconditional in-group commitment, potentially translating into extreme behaviors in a limited number of individuals who are willing to make sacrifices for their groups no matter the personal costs and consequences for others.

Current concepts of the neurobiology of altruism hold that it is anchored in the motivational architecture of the social brain. Over the past decade, research from various fields has documented the effects of the evolutionarily conserved hypothalamic peptide hormone oxytocin, in modulating a diverse repertoire of human (pro)social behaviors (Donaldson & Young 2008), in addition to its well-established role in reproduction and the formation and maintenance of pair-bonding (Hurlemann & Scheele 2016; Insel & Young 2001; Young & Wang 2004). The peptide's effects have been substantiated by a plethora of studies showing that the exogenous delivery of oxytocin as a nasal spray alters outcome measures of behavioral and neural response (Meyer-Lindenberg et al. 2011; Striepens et al. 2011; Yamasue et al. 2012). Studies have consistently shown that nasally administered oxytocin dampens amygdala reactivity toward social fear signals in both humans (Eckstein et al. 2015; Kanat et al. 2015) and macaques (Liu et al. 2015). Specifically, its potential to attenuate hypothalamic-pituitary-adrenal axis activity in response to social stressors, at least under conditions of social support (Heinrichs et al. 2003), has substantiated oxytocin's anxiolytic effect profile, which facilitates interpersonal trust (Baumgartner et al. 2008; Kosfeld et al. 2005) and the formation and maintenance of interpersonal bonds (Insel 1997).

In recent years, studies in the field of social neuroscience have provided new insights into the psychobiological mechanisms of human altruism (Hurlemann & Marsh 2016; 2017). In particular, oxytocin has been found to modulate philanthropic altruism (Israel et al. 2012; Marsh et al. 2015), empathy (Hurlemann et al. 2010; Strang et al. 2017), and parochialism and fairness (De Dreu et al. 2010; Stallen et al. 2018) along with social norm compliance (Stallen et al. 2012). The existing evidence indicates that the peptide does not invariably promote positive social behaviors but may also evoke protective-aggressive responses (Striepens et al. 2012). These effects may result from changes in perceptions of and attitudes toward the environment, including increased attention to socially relevant cues (Guastella et al. 2008) and altered sensing of and responding to emotional stimuli (Spengler et al. 2017). This resonates with evidence suggesting that the subtle, modulatory effects of oxytocin are highly susceptible to individual personality and contextual variables (Hurlemann 2017), such as gender (Scheele et al. 2014), the presence of social cues (Marsh et al. 2015), or group affiliation (De Dreu et al. 2010). In contrast to previous experimental evidence, which emphasized either the efficacy of social norms as a potential means of stabilizing altruistic cooperation (Fehr & Fischbacher 2004) or the influence of oxytocin signaling on social conformity (Huang et al. 2015; Stallen et al. 2012), a recent study combined both interventions. Therein, in a series of experiments, it was shown that oxytocin-enforced norm compliance promotes an altruistic response bias toward outsiders, even in those individuals who made more selfish decisions in the absence of such exogenous triggers (Marsh et al. 2017).

Collectively, this evidence suggests that oxytocin is centrally involved in modulating altruistic behavior, with recent studies stressing the substantial influence of message frames (Marsh et al. 2015; 2017) in determining the direction and magnitude of oxytocin effects on behavioral readouts in humans. Decomposing the underlying causes of conflict requires consideration of the evolutionary, developmental, and neurobiological origins of altruistic motivation and how individual variation with respect to this motivation may result in extreme social behaviors and actions that are dissociated from reasonable risks and rewards. Given this empirical background, we propose that pathological altruism and its determining factors may spiral into conflict, long-term hostilities or even warfare between groups, and that this behavior likely depends on and is influenced by oxytocin signaling in the brain.

# Toward the need to discriminate types of attackers and defenders in intergroup conflicts

Dashalini Katna and Bobby K. Cheon

School of Social Sciences (Psychology), Nanyang Technological University, Singapore 637332.
h160007@ntu.edu.sg bkcheon@ntu.edu.sg

doi:10.1017/S0140525X19000839, e127

**Abstract**

Here, we argue that attackers in intergroup conflicts are also likely to hold strong identity fusion, anticipate threat from the out-groups, and retaliate by signaling preemptive aggressiveness, which may not be asymmetrically exclusive to defenders. We propose that the study of the intergroup and intragroup dynamics could highlight more specific, robust markers to differentiate types of defenders from attackers.

The changing nature and complexity of conflict structures underscore the importance of interpreting attackers' and defenders' strategic positions and incentives in modern security dilemmas. The mindsets and strategies that De Dreu and Gross (D&G) suggested were largely favored by defenders in asymmetric intergroup conflicts could also play an important role among attackers. Based on identity fusion, retaliatory thinking, and intergroup threat sensitivity, we propose that, like defenders, attackers could (1) be strongly fused with their in-group; (2) be motivated to intentionally signal aggressiveness, which may especially be reflected in intimidation; and (3) engage in hostile attributions of the out-group. In particular, we suggest that these characteristics may be especially

reflective of attackers with strong tendencies toward self-sacrificial forms of prosociality and retaliatory mindsets. For these individuals, the boundaries between attack and defense may be highly malleable, with defensiveness readily motivating preemptive attacks, suggesting the need for identifying more specific, robust markers that discern their motives, strategies, and actions.

Though defender groups demonstrated stronger in-group identification (De Dreu et al. 2016a), identity fusion could also characterize attacker groups. Unlike depersonalization, which emphasizes the relegation of the personal self, identity fusion involves the merging of personal and social selves, engenders familial feelings of oneness with the ingroup, and motivates members to "do as much for the group as they would do for themselves" (Swann et al. 2012, pp. 442). Identity fusion in attacker groups may galvanize highly committed members to undertake overtly confrontational and aggressive actions by empowering their personal, agentic selves that are channeled into approach-motivated states and physical engagements directed toward the out-groups (McGregor et al. 2015). For example, far-right adherents who attach tight kinship bonds to their groups (Baray et al. 2009) directly and openly display their hostility and strength toward out-groups by staging large-scale protests that are marked by provocative racist chanting and that may culminate in physical assaults (Meadowcroft & Morrow 2017). Importantly, these types of intimidation tactics adopted to intentionally signal aggressive intent, formidability, and unity with other group members also suggest that aggressors may not necessarily seek to deceive or mismatch the target's level of defense (De Dreu et al. 2016a).

Intense feelings of fusion with the ingroup could also foster unusually strong sensitivity to detect potential threats and a large risk appetite for self-sacrificial action tendencies (Whitehouse et al. 2014). We define the tendency to voluntarily accept extreme personal costs for the welfare of other in-group members (Barclay & Van Vugt 2015; Bélanger et al. 2014) as "self-sacrificing prosociality" (Böhm 2016, p. 1). Members with high identity fusion and self-sacrificing prosociality may volunteer to undertake attack-oriented actions that involve significant risks for themselves like suicide bombers who sacrifice willingly to harm the out-group and may not serve to strategically defend the in-group (Mroszczyk 2016).

Individuals with retaliatory mindsets who seek to return potential harm with harm and rejoice when the perpetrator(s) suffers (Gerber & Jackson 2013) may tend to construe the ambiguous actions of others as hostile and aggress against the provocateur(s) readily (Topalli & O'Neal 2003). In an intergroup setting, self-sacrificial group members with retaliatory mindsets may engage in a pattern of hypervigilance (Becker et al. 2011) and elevated hostile attribution bias (Bondü & Richter 2016). These members may perceive the out-group as being deliberately inimical to the in-group (Cottrell & Neuberg 2005), and thus aggressive retaliation is appropriate and justified (Lopes & Jaspal 2015; Raihani & Bell 2018). Described as the tendency to anticipate and ascribe hostile meaning to others' behaviors in ambiguous situations (Epps & Kendall 1995), hostile attribution bias has been documented among road rage drivers (Sharkin 2004), intimate partners in conflict (Clements & Schumacher 2010), and political party supporters (Shaver et al. 2011; Waytz et al. 2014) who may motivate aggressive postures more than defensive positions.

With their in-built vigilance to detect threat for the ingroup, highly self-sacrificing members with retaliatory thinking could be most eager to signal intent to harm an out-group, relative to purely defending the in-group, which may not guarantee the safety of the in-group from the out-group (Neuberg & Schaller 2016). They may reflexively interpret any action from the out-group as antagonistic and threatening for the in-group and, in consequence, react aggressively to cues that connote possible hostility from the out-group. These members may communicate clear and advance warnings of aggression to induce fear perceptions in the out-group about its (in)capability to withstand the attack from the ingroup (Altman 2008) or loss of control (Hasan-Aslih et al. 2018). The in-group could be more fixated on deriving joy from seeing the out-group in a state of panic and confusion (Jackson et al. 2019) than deterring future attacks for the in-group as the central goal (Osgood 2017).

Consistent with the "mental hazard-precaution system" (D&G, sect. 4.2, last para.), group members with retaliatory mindsets and anticipations of conflict with the out-group may adopt an attacker posture that is prepared to self-sacrifice to inflict damage on the out-group rather than that of a defender who strives to primarily protect the well-being of the self and their in-group. Initially relying on behavioral inhibition that calls on "vigilant scanning" of threats (D&G, abstract), these members might be mobilized into behavioral activation, which potentiates them as attackers to preemptively aggress against the out-group instead of solely primarily enhancing the well-being of the in-group.

In sum, the qualities of defenders such as in-group identification, intentional signaling of aggression, and anticipating out-group hostility could also correspond to attackers in intergroup conflicts, particularly with strong self-sacrificial tendencies and retaliatory mindsets. The role of such factors indicates the need to identify interindividual differences reflecting different types of attackers and defenders, which may enable appropriate risk assessments and interventions to be developed.

# Matching pennies games as asymmetric models of conflict

Michał Wiktor Krawczyk

Faculty of Economic Sciences, University of Warsaw, 00-241 Warsaw, Poland.
mkrawczyk@wne.uw.edu.pl http://coin.wne.uw.edu.pl/mkrawczyk/

doi:10.1017/S0140525X19000864, e128

**Abstract**

De Dreu and Gross (D&G) seem to have disregarded some relevant experimental literature on games of conflict, most notably variations on "matching pennies" games. While in such games, "attacker" and "defender" are typically not explicitly labelled, players' differentiated roles yield naturally to such notions. These studies partly validate some of D&G's findings and interpretations.

In their engrossing article, De Dreu and Gross might have overlooked some directly relevant literature on asymmetric games of conflict. In particular, they largely disregard the literature on (generalised) matching pennies (MP) games (only citing Goeree et al. 2003 in passing). In section 2.1, especially Figure 1, they appear to define the binary Attacker-Defender Game (AD-G) as a combination of the assurance game and the game of Chicken; this leads them to keep identical labels of cooperation and defection for both players'

**Table 1.** (Krawczyk) De Dreu and Gross's AD-G (A) and Goeree et al.'s Game 4 (scaled down and with rows reversed to facilitate a comparison), a generalised MP game (B)

| A | | Attacker | |
|---|---|---|---|
| | | C | D (seek gain) |
| Defender | C | 2,1 | 0,2 |
| | D (protect) | 1,1 | 1,0 |

| B | Left | Right |
|---|---|---|
| Up | 37, 20 | 1, 37 |
| Down | 20, 16 | 16, 1 |

strategies in AD-G, which is debatable given that the game is asymmetric. Much more importantly, assurance and Chicken games may not be the nearest relatives of the binary AD-G that have been considered in the literature; this AD-G is just an example of the well-known class of generalised MP games (also referred to as *inspection games* or *police-public games*). In either case, one of the players wants to match and the other wants to mismatch (see Table 1) so that the only equilibrium is one in mixed strategies, which is, however, Pareto inefficient. The main difference would be in the framing (which may matter in such games; see Eliaz & Rubinstein 2011). Indeed, the labels of "attacker"/"defender" are not typically used in (generalised) MP games and, in some of these games, the strategies could not reasonably be labelled "cooperation" and "defection," because when the "defender" is "protected," it may actually benefit from the "attacker" "seeking gain"/"attacking."

Arguably, each of the three possibilities (the defender being indifferent when protected, as in Table 1A; the defender being hurt by the attack even when protected, although not as much as when unprotected, as in Table 1B; the defender benefitting from the attack when protected, as in standard MP games) may be the most appropriate, depending on the conflict situation at hand. For example, the cost of war will usually still be serious for the defender, even if it is well prepared and ultimately victorious; on the other hand, such an episode may strengthen its international position as a powerful player and scare other potential aggressors away, and so on. Similar consideration applies to the attacker's preferences concerning the defender's actions when the attacker does not seek gain.

A number of findings from MP games parallel those reported by D&G. In particular, players tend to react in natural ways to their opponent's probability of choosing each option in the past (e.g., Colman 1999). Moreover, in Goeree et al.'s Game 4, shown here in Table 1B, "attackers" (column players) choose the "defective" strategy Right (hurting the other player) significantly less often than the Nash equilibrium would require, but there is no analogous effect for the "defenders." This is equivalent to D&G's findings in continuous AD-Gs. Then again, the evidence for "attackers" being relatively timid is rather mixed in Dorris and Glimcher (2004), Rauhut (2009), and Nosenzo et al. (2013).

D&G's claim that "theory and research has rarely made a clear distinction between attack and defense" (sect. 1, para. 4) is also potentially slightly misleading when being applied to models of conflict with *continuous* action space. Indeed, several papers have analysed asymmetric conflict situations both theoretically (Franke et al. 2013; Nti 1999) and empirically (Carter & Anderton 2001; Dechenaux et al. 2015, p. 623). One frequently considered type of heterogeneity is that of a player possessing more resources and trying to keep them whilst another seeks their redistribution, which may be naturally interpreted as a defence-attack situation. Further, in "multi-battle" contexts, the goals of the defender and the attacker are often explicitly differentiated (Deck & Sheremeta 2012; Kovenock et al. 2010).

Finally, there is literature looking at process data (such as reaction times, eye-tracking, and neuroimaging) in MPs (Hampton et al. 2008; Krol & Krol 2017). While these typically involve the standard zero-sum MP game, still the player who tries to match may be thought of as a defender, while the one who tries to mismatch may naturally be construed as an attacker. For example, D&G's proposition that defenders act more spontaneously than attackers is preceded by the data of Martin et al. (2014), in whose experiments the matchers were faster than the mismatchers. These authors interpret it in terms of humans' automatic tendency to imitate (Belot et al. 2013). To recapitulate, a more complete survey of the relevant extant literature, even if it involves different labels in analogous games, could provide insights into the robustness of D&G's reported findings and their interpretations.

# The importance of raiding ecology and sex differences in offensive and defensive warfare

Anthony C. Lopez

School of Politics, Philosophy and Public Affairs, Washington State University, Vancouver, WA 98686.
Anthony.c.lopez@wsu.edu https://labs.wsu.edu/evolutionary-politics/

doi:10.1017/S0140525X19000906, e129

**Abstract**

De Dreu and Gross offer a compelling synthesis of a growing literature on the psychology of attack and defense. I argue that human raiding ecology suggests the need to endogenize attacker-defender move order as well as opportunities for tactical mismatch available to defenders. Perhaps most significantly, I draw attention to the surprising lacunae in sex differences across attack and defense.

De Dreu and Gross (D&G) deliver a compelling case for the existence of separate psychologies of attack and defense, and there are ample opportunities for caveats and extensions in this productive line of inquiry. I note three, in particular, and discuss each in turn. First, an ecologically valid model of attack versus defense should acknowledge an asymmetric move order between the two. Second, while defenders seek to match attackers strategically, they may simultaneously seek to mismatch them tactically. Finally, we should expect sex differences to be particularly significant in offensive but not defensive warfare. I restrict my discussion largely to coalitional as opposed to individual attack and defense.

## *Attackers move first*

The most common form of intergroup violence in evolutionary history has been the raid, which, in part, is defined by its move

order asymmetry. Specifically, attackers move first. When attackers move first, they are likely to choose D when: (1) defenders have *already* chosen C (however unwittingly; e.g., they are asleep, facilitating an ambush) or (2) defenders are weak. Similarly, defenders are unlikely to choose D unless attackers have *already* chosen D, because, as the authors say: "Unilateral defense is costly." The above implies that attackers will choose D in a first move when: (1) the risks are sufficiently low (Lopez 2016; Tooby & Cosmides 1988; Wrangham 1999) or (2) social institutions provide countervailing incentives to sustain coordinated behavior (Glowacki & Wrangham 2013). Importantly, endogenizing move order is not only consistent with the analysis provided by D&G but also it is potentially more reflective of ancestral landscapes of intergroup violence (Gat 2006; Wrangham & Glowacki 2012).

### Defenders can run, absorb, and retaliate

A value of the Attacker-Defender Game is its simplicity, which allows us to ask questions about what may be missing, whether what is missing is important, and how to integrate these elements. For example, one can see that both "defending" and "not defending" can have either active or passive manifestations. "Defending" can mean active physical denial, or it can mean a relatively passive sunk cost investment in defensive capabilities (e.g., walls & weapons). "Not defending" can mean anything from surrender, to being caught unaware, to running away. Indeed, the latter was likely a common and prudent choice in ancestral environments, although it is obviously and tellingly infeasible for modern nation states. In short, varieties of "not defending" may carry different payoffs; surrendering potentially means death, while running away allows, at a minimum, the possibility of survival, and at a maximum, opens the possibility of reorganization and retaliation. Thus, "running away to fight again another day" may also be properly understood as asymmetric defense, which suggests that defenders may profit just as much from cultivating asymmetries as their adversaries.

Specifically, under certain conditions, the defender preference for matching attacker *strategy* may be complemented by a defender preference for mismatching attacker *tactics*. One of these conditions is likely to be when defenders are weak. When attackers choose D, even weak defenders may prefer D > C when they have the option to run from or absorb an attack and retaliate. International relations scholars note that weaker adversaries are often able to win wars precisely when they can *mismatch* the conventional tactics of their attackers with unconventional guerilla warfare (Arreguin-Toft 2005). Large nation states exploit a similar principle – the notion of "elastic defenses" – when they absorb an attack while preparing a counterattack, as the Russian Empire did to great effect against the advancing armies of Napoleonic France (Biddle 2006). Thus, Wrangham and Glowacki (2012) insightfully note that one of the elements that makes human coalitional violence so much riskier than chimpanzee violence is the likelihood of retaliation and revenge, which again works to undermine attacker success. This is particularly true in combination with the problems of coordination and motivation that D&G usefully discuss.

### Sex differences exist

The authors present compelling evidence that defense benefits from rapid and spontaneous in-group support, while attack is vulnerable to failures of motivation and coordination. One additional wellspring of motivation the authors do not consider is sex differences in aggression (Daly & Wilson 1988; Goldstein 2003; McDonald et al. 2012; Van Vugt 2009). A growing body of evidence suggests that males and females likely possess distinct *conditional* psychologies of aggression (Brooks & Valentino 2011; Ginges & Atran 2011; McDermott 2015), and it is likely that these psychologies find unique expression in attack and defense.

One of the key features distinguishing attack from defense is that the latter is more clearly a public goods problem (Lopez 2010; Rusch 2013; 2014a; Tooby & Cosmides 1988). This fact alone suggests that, all things equal, there should be no sex difference in support for investing in defense. What it does not suggest, however, is that females should be just as willing as males to *physically* fight. Indeed, experimental evidence indicates that, although males and females anticipate similar levels of benefit from successful defense, males profess a greater willingness to directly participate in the fighting. In contrast, males report greater anticipated benefit from *and* a greater willingness to participate in offensive coalitional aggression relative to females (Lopez 2017). In other words, although males are more physically aggressive across offensive and defensive scenarios, females are more likely to mirror male hawkishness in defense than in offense. Relatedly, Wrangham notes that "proactive aggression has been predicted to be associated with [larger] sex differences, consistent with the greater prevalence of psychopathy among males" (2018, p. 250). Taken together, we should expect sex differences in aggression to reflect an adaptive logic that functionally embodies the distinct challenges of offense and defense, which is consistent with both the "male warrior hypothesis" (McDonald et al. 2012) and with observations of greater proactive aggression and psychopathy among males relative to females.

The argument for separate psychologies of attack and defense that D&G lay out is compelling; provides a useful example of the functional link between evolution, psychology, and behavior; and affords many lucrative opportunities for future study.

## Moral rigidity as a proximate facilitator of group cohesion and combativeness

Antoine Marie

Institut Jean Nicod, Ecole Normale Supérieure, PSL Research University, 75005 Paris, France.
antoine.marie.sci@gmail.com
https://antoinemariesci.wixsite.com/antoinemarie

doi:10.1017/S0140525X1900092X, e130

**Abstract**

De Dreu and Gross's description of the proximate mechanisms conditioning success in intergroup conflict omits humans' deontological morality. Drawing on research on sacralization and moral objectivism, I show how "moral rigidity" may have evolved through partner selection mechanisms to foster coalitions' cohesion and combativeness in intergroup conflict.

De Dreu and Gross's argument that attack and defense are distinct strategies underpinned by different neuropsychological circuitries is an original refinement of the theory of conflict.

However, their description of the proximate mechanisms facilitating success in intergroup competition (sect. 4, target article) omits humans' deontological moral intuitions. In interaction with overconfidence biases, hostile attributions, and the enforcement of "cultural rituals and sanctioning systems" (sect. 4, para. 1), what may crucially help groups of individuals cohere and prevail in conflict are high levels of "moral rigidity" in their tribal members, that is, of inflexible loyalty to their interpersonal commitments within the group.

## Proximate processes of moral rigidity

*Sacralization* is the valuation of core social obligations (e.g., loyalty to comrades, individual freedom), symbols (the Koran, the flag), or resources (the Holy City, Hindus' cows) far beyond their practical utility. Individuals who sacralize tend to disregard consequences, opportunity costs, and the maximization of aggregate welfare, and to perceive trade-offs with secular values as taboo and morally contaminating (Atran 2010; Baron & Spranca 1997; Graham & Haidt 2012; Tetlock 2003; Tetlock et al. 2000). *Folk moral objectivism* is the intuitive propensity to represent some moral obligations as "facts" of nature, externally imposed on human wills and irreducible to mere subjective preferences (Beebe et al. 2015; Goodwin & Darley 2010; 2012; Stanford 2018).

Sacralization as a motivational process, and folk objectivism as an intuitive ontological commitment (Quine 1948), are low-level cognitive mechanisms. Both tendencies underlie what I propose to call "moral rigidity," the inclination to treat key interpersonal obligations as non-negotiable and to reify them.

Individuals and subcultures vary as to which moral norms, "foundations," and resources they sacralize (Graham et al. 2013; Haidt 2012; Atran 2010). Moreover, recent studies on Western samples found variability in the degree to which individuals provide deontological (vs. consequentialist) justifications for their moral judgments (Piazza & Sousa 2013) and objectify moral prescriptions (Goodwin & Darley 2010; 2012). Still, evidence suggests that children are predisposed to moral objectivism (Nichols & Folds-Bennett 2003; Wainryb et al. 2004), and that most normally developing individuals, regardless of political orientation, have some sacred values (Frimer et al. 2017; Haidt 2012). Taken together, advances support the hypothesis that human minds may be innately prepared for moral rigidity, understood as a set of domain-specific, yet culturally flexible, cognitive adaptations to the demands of social life.

Why would the cognitive adaptations underlying moral rigidity have evolved? And what role do they play in intergroup conflict? I defend the hypothesis that humans evolved tendencies to sacralize and reify moral obligations so as to behave and be seen as devoted partners, and avoid threats on their coalition's cooperative and competitive potential. These tendencies, which proximately manifest as moral rigidity, would have helped maintain costly investment in risky coalitionary ventures, from collaborative hunting to violent intergroup conflict.

### *From interdependence to a proportionality-based morality*

Much of hominid life has been about achieving social integration in fluid groups by (1) reliably signaling one's willingness to respect others' welfare, while simultaneously (2) protecting oneself from exploitation. Evolutionary theorists (Alexander 1987; Baumard et al. 2013; Frank 1988; Trivers 1971) have argued that selection pressures resulting from competitive altruism (Barclay & Willer 2007), and increasingly refined abilities to sanction insufficient dedication in potential partners (e.g., gossip and shunning), have selected for moral preferences calibrated for securing conditional cooperation (in hunting, gathering, shared parenting, combat, etc.), by sharing its costs and benefits equally among stakeholders. As a result, humans evolved intuitions that individuals with whom they interact are persons with inalienable rights, and that one's commitment to their interests should be proportionate to the amount of benefits one can expect to secure (or losses one can hope to avoid) through them (Aktipis et al. 2018; Baumard et al. 2013).

### *Moral rigidity as an error management "bias"*

The logic of partner choice mutualism implies that to any type of joint venture corresponds a minimal level of cooperative engagement that each team member would obligatorily demand of his partners. As long as (i) groups cohabitated peacefully or natural resources abounded, for instance, one can expect within-group interdependence to have been moderate. Social selection pressures would have mainly been on individuals' ability to demonstrate unconditional respect for their in-groups' physical integrity and property, and the community's resources, to avoid being left out of everyday ventures like hunting and gathering. However, (ii) ancestral hominins also faced recurrent intergroup competition and warfare (LeBlanc & Register 2003; Tooby & Cosmides 2010). In such circumstances, lethal threats coming from outside the group and the possibility of losing reproductive resources (such as territory or women) to rival coalitions would have driven costs of insufficient commitment to skyrocket, thereby dramatically elevating the level of prosocial engagement demanded of each group member.

Insufficient dedication to one's partners' interests, whether in situation (i) or (ii), would on average have led to reproductive impasses. In terms of error management theory (Haselton et al. 2015), evolution should therefore have selected for moral heuristics that are "biased" in favor of the most adaptive strategy: (1) making absolutely sure the individual will display the minimal level of prosocial commitment demanded by the contextual level of interdependence, while (2) showing automatic aversion to acts that could endanger the coalition's cooperative potential and combativeness in intergroup conflict. In this respect, intuitively sacralizing and objectifying core interpersonal obligations may constitute the most cost-effective computational way of completely isolating them from the temptation to trade them off with immediate self-interest (Marie & Fitouchi, in preparation).

In line with this hypothesis, experiments find that potential partners expressing deontological moral judgments in everyday settings – a proxy of underlying moral rigidity – are judged more trustworthy and are more likely to be chosen (Everett et al. 2016). Furthermore, the anthropology and psychology of contemporary parochial altruism in military, insurgent, and terrorist behavior suggest that moral rigidity may have been shaped by ancestral warfare to bolster defense of in-groups, territory, and communal resources. The values and goals for which contemporary attackers and defenders fight (e.g., the Reich, the Caliphate) often undergo ritualistic processes of sacralization and objectification (Atran 2010; 2016), up to a point where activists end up

representing them as absolute coordinates of reality, and become blinded to exit strategies. By turning fighters into "devoted actors," activation of high levels of moral rigidity – often complemented by "identity fusion" (Whitehouse 2018) – predicts individuals' willingness to engage in costly sacrifice, including death, for their comrades and cause and *ceteris paribus* significantly increases fighters' chances to prevail in combat (Atran 2016; Gómez et al. 2017).

**Acknowledgments.** I thank Léo Fitouchi, to whom my thought on "moral rigidity" is heavily indebted, and Manvir Singh, for his useful remarks on earlier drafts.

# But how does it develop? Adopting a sociocultural lens to the development of intergroup bias among children

Niamh McLoughlin and Kathleen H. Corriveau

Wheelock College of Education and Human Development, Boston University, MA 02215.
niamhmc@bu.edu kcorriv@bu.edu
http://www.bu.edu/wheelock/profile/niamh-mcloughlin/
https://www.bu.edu/wheelock/profile/kathleen-corriveau/

doi:10.1017/S0140525X19000761, e131

**Abstract**

We argue that adopting a sociocultural lens to the origins of intergroup bias is important for understanding the nature of attacking and defending behavior at a group level. We specifically propose that the potential divergence in the development of in-group affiliation and out-group derogation supports De Dreu and Gross's framework but does indicate that more emphasis on early sociocultural input is required.

Based on theory and empirical findings drawn from the fields of neuroscience, behavioral economics, social, and evolutionary psychology, De Dreu and Gross develop a framework to disentangle the neurobiological, psychological, and cultural characteristics that contribute to the nature of intergroup conflict. More specifically, they propose that potentially distinct psychological and cultural factors could be involved in what they conceptualize as "attacking" versus "defending" behavior at a group level. We agree with the argument that social affiliation, such as identification with and loyalty to in-group members, emerges more spontaneously within defender groups, whereas the formation and behavior of attacking groups may rely more heavily on cultural tools and arrangements (e.g., propaganda, cultural rituals, religious belief). However, research in developmental psychology that adopts a sociocultural lens to examine the origins of intergroup bias is critical to understanding these phenomena.

Children are motivated to affiliate with others from early in development (Carpenter et al. 1998; Over 2016) and particularly with individuals who belong to their own social group. A substantial number of studies have demonstrated that 4- and 5-year-old children are more likely to selectively befriend (Kinzler et al. 2007), stay loyal to (Misch et al. 2016), share more resources with (Renno & Shutts 2015), and trust the testimony of (Harris & Corriveau 2011) in-group members. These social affiliative tendencies are observed in societies in which intergroup tension is relatively low and even when group membership is based on very minimal cues (i.e., shirt color; Dunham et al. 2011). De Dreu & Gross posit that the spontaneously coordinated and affiliative behavior witnessed within defense groups could be due, in part, to the earlier emergence of defense behavior in ancient societies. The early appearance of affiliative behavior in development also supports the evolutionary value of belonging to a group (Brewer 1999; Brewer & Caporael 2006). Although the drive to affiliate with socially similar others is considered to be a fairly ubiquitous phenomenon (Meltzoff 2007), some researchers have emphasized the role of early sociocultural experiences in the development of in-group positivity. For example, minority race children can exhibit less favoritism for in-group members compared with majority (white) children (Margie et al. 2005; Shutts et al. 2011), and variability in socialization with and exposure to different racial groups can shape young children's social preferences (Chen et al. 2018; Gaither et al. 2014).

Other research in developmental psychology has suggested that negative out-group attitudes appear slightly later in development, typically after the age of 6 (Buttelmann & Böhm 2014; McLoughlin & Over 2018). A recent body of work has illustrated that older children (ages between 6 and 11 years) can even hold dehumanizing out-group views, including the belief that out-group members possess fewer mental state capabilities (Dore et al. 2014; McLoughlin & Over 2017) and uniquely human characteristics (Chas et al. 2018; Costello & Hodson 2014). Yet, there is evidence to suggest that the developmental onset and trajectory of harmful out-group perceptions can vary widely, depending on different sociocultural factors, such as the cultural saliency of historical conflict between ethnic groups (Bennett et al. 2004; Birnbaum et al. 2010; Smyth et al. 2017), as well as the shared ideologies of the surrounding community (Rhodes & Mandalaywala 2017; Segall et al. 2015). Taken together, these findings reinforce the importance of cultural learning and communication (Harris et al. 2018; Over & McCall 2018; Skinner & Meltzoff 2019) and further speak to the relevance of social tools in motivating attacking behavior.

In sum, the potential divergence in the development of affiliation with the in-group and derogation of the out-group – at least in the cultural contexts studied thus far – supports the authors' proposed framework but does indicate that more emphasis on early sociocultural input is required. One caveat to consider, however, is that it is sometimes not feasible to experimentally manipulate or measure parallel behaviors in research with children (e.g., extreme attacking intentions). Nonetheless, this field can provide valuable insight into the adaptive psychological mechanisms and the social learning processes contributing to intergroup conflict.

# Behavioural inhibition and valuation of gain/loss are neurally distinct from approach/withdrawal

Neil McNaughton[a] and Philip J. Corr[b]

[a]Department of Psychology, University of Otago, Dunedin 9054, New Zealand and [b]Department of Psychology, City, University of London, London EC1R 0JD, United Kingdom.

neil.mcnaughton@otago.ac.nz Philip.Corr.1@city.ac.uk
http://www.otago.ac.nz/psychology/staff/neilmcnaughton.html
https://www.city.ac.uk/people/academics/philip-corr

doi:10.1017/S0140525X19000712, e132

**Abstract**

Gain or omission/termination of loss produces approach; while loss or omission/termination of gain produces withdrawal. Control of approach/withdrawal *motivation* is distinct from *valuation* of gain/loss and does not entail learning – making "reward" and "punishment" ambiguous. Approach-withdrawal goal conflict engages a neurally distinct Behavioural Inhibition System, which controls "anxiety" (conflict/passive avoidance) but not "fear" (withdrawal/active avoidance).

In section 3.1, De Dreu and Gross contrast reward seeking with loss aversion and conflate behavioural inhibition with fear and active avoidance. We argue that this confuses: (1) valuation with motivation; (2) anxiety with fear; and (3) reinforcers with reinforcement. Making these distinctions has consequences for their proposed neuropsychology.

The expectation/availability (innate or learned) of gain elicits approach. However, omission/termination of expected gain elicits defensive withdrawal (Adelman & Maatsch 1956) and attack (Gallup 1965; Kelly & Hake 1970), as does an *explicit* aversive stimulus, such as shock (Renfrew & Hutchinson 1983). Importantly, even in the presence of loss aversion (Kahneman & Tversky 1979; Novemsky & Kahneman 2005; Tversky & Kahneman 1991), approach tendencies can be stronger than withdrawal (Hall et al. 2011), likely as a result of their different goal gradients (Brown 1948). So, approach/withdrawal motivations are controlled independently of gain/loss valuations (Hall et al. 2011), and it is important to keep valuation and motivation theoretically separate (Corr & McNaughton 2012) and always take into account the role of contingency (Fig. 1).

It is also important to keep "anxiety" separate from "fear." Despite their frequent semantic conflation (McNaughton 2018), the neuropsychology and psychometric evidence are clear on their differences (Corr et al. 2013). In contrast to a fear/withdrawal system that is sensitive to threat, the anxiolytic-sensitive Behavioural Inhibition System (Gray 1977) processes goal-conflict and amplifies behavioural inhibition/passive avoidance/defensive quiescence, attention, arousal, and negative bias (Gray & McNaughton 2000; McNaughton & Corr 2004). This is neurally distinct from the panicolytic-sensitive systems that mediate fight, flight, freezing, and withdrawal/active avoidance (Fig. 2), collectively known as the Fight-Flight-Freeze System (FFFS). Note that "fight" in this context is a defensive response and quite distinct as a behaviour from the predatory "attack" that is contrasted with "defense" in the target article – although in human personality questionnaire studies the relations between withdrawal, defensive attack, and predation are unclear (Corr 2016). Contrary to the picture painted by De Dreu and Gross, it is anxiety rather than fear that is linked to the release of stress hormones like cortisol (see McNaughton 1989, pp. 57–59); and, while $5HT_{1A}$ agonists are anxiolytic but not panicolytic, the serotonin system as a whole innervates and affects not only the Behavioural Inhibition System but also the withdrawal and the approach systems, with quite high-level consequences (Carver et al. 2008).

For the same reasons, we think their picture of prefrontal control networks should be split and extended to subcortex. We agree that anxiety involves the inferior frontal gyrus (Shadli et al. 2015), basolateral amygdala, hippocampus, ventromedial orbital cortex (Fig. 2), and insula (Paulus & Stein 2006). However, we would add the posterior cingulate cortex, and, with risk assessment in particular (McNaughton & Corr 2018), there is an important role for subcortical "survival circuits" (Mobbs & LeDoux 2018) that include the dorso-lateral and ventro-lateral periaqueductal grey, anterior and lateral hypothalamus, and lateral septum (Motta et al. 2017). Critically, we see fear as neurally distinct, involving lateral orbital cortex, anterior cingulate, central amygdala, medial hypothalamus, and dorso-medial periaqueductal grey.

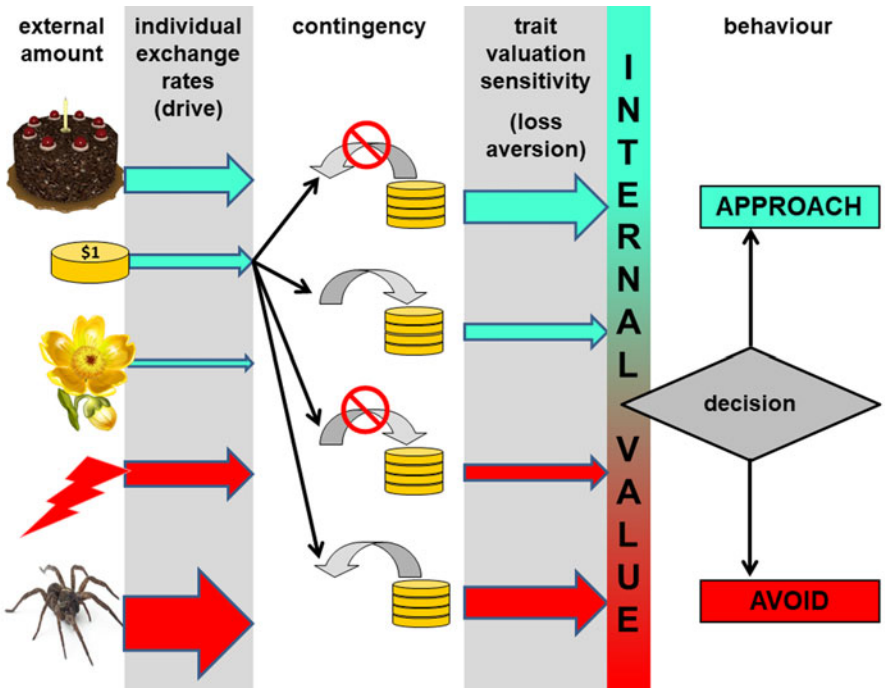


**Figure 1.** (McNaughton & Corr) Relations between external amount, contingency, and value. An external item will have a specific amount (e.g., one entire cake) that, together with the current level of drive (which acts like a currency exchange rate) for that kind of item for that person, determines its primary internal value (thickness of *arrows* in first column). As shown for the case of $1, this interacts with whether the item will be gained or lost to determine the direction and size of its internal value, as ultimately measured by the effect on behaviour. The direction of this effect is reversed if the [expected] gain or loss is omitted. Loss (removal from a store of items) is most easily controlled with money but will also occur when, for example, one rat steals the food from another rat. Reprinted, with permission, from McNaughton et al. (2016).

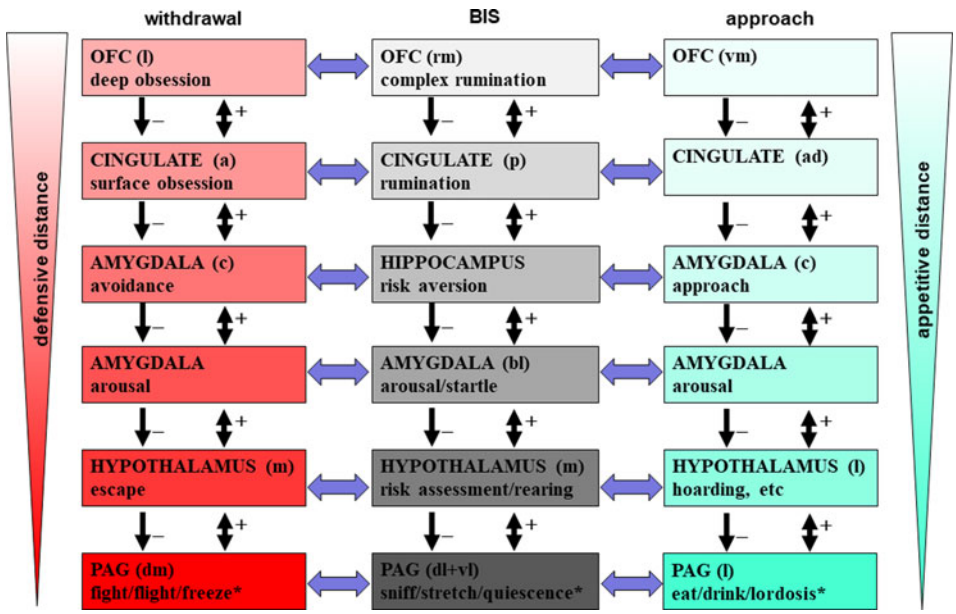


**Figure 2.** (McNaughton & Corr) Hierarchical organization of approach, withdrawal, and behavioural inhibition (BIS) in terms of behaviour and neural level. Lower levels process small defensive distances; higher levels process greater ones (i.e., negative events that are more distant in space or time). Activation tends to spread through the whole system (*double-headed black arrows*), but strong activation of a higher level (e.g., avoidance) inhibits (*single-headed arrows*) the behavioural output from (but not the activation of) lower levels (e.g., escape). *Static postures that achieve withdrawal, conflict resolution, or approach, respectively. a = anterior; ad = anterodorsal; bl = basolateral; c = central; dl = dorsolateral; dm = dorsomedial; l = lateral; m = medial; OFC = orbital frontal cortex; p = posterior; PAG = periaqueductal grey; rm = rostromedial; vl = ventrolateral; vm = ventromedial. Reprinted, with permission, from McNaughton and Corr (2018).

The three-system (approach, withdrawal, conflict) hierarchical neuropsychology we have described is also relevant to the trait considerations of section 3.6. "These systems mediate fluid moment-by-moment reactions to changing stimuli, with relatively stable person-specific sensitivities to these stimuli manifested in personality traits" (Corr et al. 2013, p. 158) and are the basis for the Reinforcement Sensitivity Theory of personality (see Corr 2008). Our perspective (avoiding the ambiguity of "reward") suggests that attack (as a predatory approach tendency) and defense (functioning to allow withdrawal) likely depend on fundamentally similar hierarchical system architectures. Apparent prefrontal versus subcortical control differences between them likely depend on the usual difference in "motivational distance" in their eliciting situations. Initially, at least, a predator will be at a large appetitive distance from the prey, requiring extensive planning of its attack. Especially where an ambush is involved, the defensive response by the prey will be immediate and even undirected (at zero defensive distance). Impulse control also involves a balance between approach motivation and inhibition. The strength of inhibition can be affected by variations in conflict sensitivity (Gray & McNaughton 2000) and in loss aversion (Tversky & Kahneman 1991); and approach can vary with the strength of delay discounting (Frost & McNaughton 2017). The effects of traits on attack and defense clearly require a highly nuanced approach.

# Towards the elucidation of evolution of out-group aggression

Nobuhiro Mifune[a] and Dora Simunovic[b]

[a]School of Economics & Management, Kochi University of Technology, Eikokuji, Kochi City, Kochi 780-8515, Japan and [b]Field C, Bremen International Graduate School of Social Sciences, 28759 Bremen, Germany.
n.mifune@gmail.com doras@bigsss-bremen.de

doi:10.1017/S0140525X19000943, e133

**Abstract**

We focus on the implications of De Dreu and Gross's findings for the evolutionary perspective on out-group aggression and in-group cooperation. Although their experimental protocols are potentially useful in determining the origins of out-group aggression in humans, they so far provide inconclusive evidence only. We suggest ways of furthering our understanding of the connection between parochial cooperation and intergroup conflict.

The mystery of human intergroup aggression has been studied in various fields of social sciences, from anthropology to political sciences. Previous psychological research mainly focused on attacking behaviour, leaving the psychology of defence considerably neglected. Although experimental economic games can be used to separately measure attack- and defence-based interactions between multiple individuals, research has so far not focused on the asymmetry that results when attack and defence co-occur, as they most often do. In this sense, the Attacker-Defender Game (AD-G) and IAD-C offer significant methodological and theoretical developments. In this commentary, we mainly focus on how these games can be used to uncover the mystery of the evolution of out-group aggression. We discuss their future directions, as well as their limitations.

Modern investigations of intergroup behaviour rely on integrating theories from evolutionary sciences with economic games. Particular focus is given to the human tendency to cooperate with in-group members and display aggression towards out-group members. Here, we will focus on the co-evolution model (CM; Choi & Bowles 2007). The CM suggests that frequent intergroup conflict in human evolutionary and cultural past caused parochial cooperation, which psychologically co-occurs with out-group aggression. In other words, at its core, the CM predicts that humans should be unconditionally motivated to not only increase their own group's standing, but also decrease the standing of other groups. Many subsequent studies attempted to test this prediction

using various experimental paradigms, from third-party punishment (Bernhard et al. 2006), intergroup prisoner's dilemma – maximizing difference (Halevy et al. 2008), and the pre-emptive strike game (Simunovic et al. 2013). While some researchers found support for the predictions of the CM, others did not (e.g., Yamagishi & Mifune 2016).

De Dreu and Gross (D&G)'s AD-G and IAD-G protocols offer new experimental paradigms to test the validity of the CM, as they may possibly and uniquely be able to appropriately reproduce the complex intergroup environment assumed by the CM in the laboratory. We discuss each game separately to illustrate what it tells us and what it can tell us about the co-evolution of out-group aggression and in-group cooperation.

In the case of IAD-C, the results reported by D&G do not provide unique support for the CM. Participants invested more in defence than in attack. Defence was more common and intuitive a choice than attack, which goes against the idea that humans are equally motivated for in-group cooperation and out-group aggression. There is some ambiguity here, however, as a non-negligible number of attacks occurred. In addition, as the authors suggest, we may suppose only defensive aggression was the mechanism that supported in-group-directed cooperation, thus making defensive, but not offensive, out-group aggression "the midwife of altruism and xenophobia" (sect. 5.1, para. 3).

AD-G, by contrast, does not at present inform new insights into the CM, but it has great potential value to do so. By introducing minimal group manipulation into the AD-G, we would be able to compare individual attack rates towards in-group and out-group members. We expect out-group members to be attacked and defended against more often than in-group members. The CM predicts no change to these attack and defence rates if the ties of interdependence between in-group members are broken (e.g., Balliet et al. 2014; Yamagishi & Mifune 2016). However, it is possible that we observe a reduction of offensive but not defensive actions in such circumstances, thus providing support for the idea that defence was the driver of the psychological adaptations responsible for in-group-bounded cooperation.

D&G also report that defence was a more intuitive decision for participants to make than was attack, suggesting that it is a default strategy in conflict situations. This is evident in both the AD-G and the IAD-C, but once again we contemplate an intergroup AD-G. Previous research may have shown that defence is not the automatic response to the presence of out-group members. Defensive aggression, measured by the pre-emptive strike game, has not emerged as either the dominant individual choice (Mifune et al. 2017) or a 3-on-3 paradigm (Mifune et al. 2016). This could be because defensive behaviour in intergroup situations occurs only when people perceive distinct attack intentions in out-group members, thus negating its unconditional nature as predicted by the CM. This should be investigated using the AD-G and with special reference to the prediction of the out-group partner's behavioural preferences.

Attacking behaviour can, at least in part, be deconstructed into preference and belief, as can be done for cooperation (Yamagishi et al. 2013). Preference, such as loss or inequity aversion, refers to one's outcome goals. Beliefs about or expectations of others' behaviours, partly driven by the perception of the ongoing situation, refer to the individual's outcome expectations (Pruitt & Kimmel 1977). D&G's results open the possibility that expectation is a more important factor in attack/defence relations than is preference. This has already been demonstrated for in-group cooperation (Balliet et al. 2014; Yamagishi & Mifune 2008). The two novel games have the potential to become the key to deciphering that puzzle while simultaneously providing special ecological validity inherent in the fact that participants would be interacting in the same paradigm, rather than in separate ones, to measure attack and defence.

In conclusion, D&G suggest methodological and theoretical improvements to the field of intergroup relations, which we find important and fruitful. One of the avenues their work opens has to do with the in-group cooperation and out-group aggression CM. By demonstrating the dichotomy and asymmetry inherent in a large proportion of intergroup conflicts, their results suggest possible refinements for the CM. In particular, the AD-G and IAD-C show that defensive rather than offensive out-group aggression could be the theoretical underpinning that would allow for the development of an in-group-bounded cooperative psychology in humans. In addition, these games provide opportunities to separate attack from defence while retaining ecological validity and to test the relative importance of outcome preference and outcome expectations. Nevertheless, more research is needed to claim anything more conclusive.

# The evolutionarily mismatched nature of modern group makeup and the proposed application of such knowledge on promoting unity among members during times of intergroup conflict

Jiaqing O

Department of Psychology, Aberystwyth University, Ceredigion SY23 3UX, United Kingdom.
ojthehuman@gmail.com
https://www.aber.ac.uk/en/psychology/staff-profiles/listing/profile/jio2

doi:10.1017/S0140525X19000797, e134

**Abstract**

Many modern-day groups differ from prehistoric ones regarding the proportion of members who are related to any particular individual. From an evolutionary mismatch lens, an appreciation of this disparity could help better explain the potential dilution of group cohesion during peacetime and inform novel, more effective approaches to enhancing group unity – strategies that might enhance national security around the globe.

De Dreu and Gross put forth a highly interesting and thought-provoking account of the art of war involving both within-group and intergroup hostilities from a scientific point of view. The authors also offer a useful overview and explanation of how within-group sense of affiliation and collective action can be potentially enhanced by capitalizing on members' sensitivity to external menaces and through using a variety of innovative strategies to ensure loyalty. However, I contend that such knowledge regarding the relationship between group identification and

cohesive responses in times of conflict, especially in the modern world, can be further (and significantly) enhanced by incorporating insights from an evolutionary mismatch perspective. Defined as the notion that many present-day issues occur as offshoots resulting from the disparity between the extremely sluggish process required for genes/psychological mechanisms to evolve and the swift changes in physical and social circumstances that have transpired since prehistory (O 2018a), evolutionary mismatch has been used as a useful framework in interpreting a diverse array of phenomena that range from zoophobia (O 2018b) to reproductive intentions (Li et al. 2015) – and could likewise be used to deepen the understanding of allegiance and cooperation among group members in conflictual situations.

First, specifically, although the authors' assertion that a lack of any foreseeable attack by foes (e.g., during peacetime) will potentially result in the disintegration of a deterrence-oriented group (because members do not need to be united to ward off any act of aggression) makes perfect sense on the face of it, such a suggested phenomenon can veritably only be understood thoroughly through the lens of an evolutionary mismatch. In particular, the evolutionary mismatch paradigm could conceivably provide a useful explanation regarding the supposedly fragile nature of group solidarity among individuals in deterrence-oriented groups. Based on such a framework, the logical question to ask is whether the potential dilution of in-group identification and cohesion in conflictual situations (especially during a lull period of relative peace) a modern-world problem.

In light of prevailing knowledge that humans have largely organized themselves together into relatively small, closely knit, kin-based groups across much of human evolutionary history (Dunbar 1993; Hill et al. 2011), the appearance of huge, broadly non-biologically (and non-affinally) related and loosely connected groups (e.g., countries) in the present day appears to be at odds with the kind of adaptive context humans have evolved from that is favorable for establishing a generally unfaltering attachment to the group. Humans are expected to be more likely to remain loyal and committed to defending a group mainly made up of biological and affinal kin (than unrelated individuals) because the well-being of these members would be directly/indirectly conducive in enabling their own genes to be passed down to the subsequent generations (Hamilton 1964a; 1964b; Hughes 1988). In virtue of these, it is reasonable to proffer that, rather than it being a natural consequence due to a lack of any possible act of aggression from external foes, as suggested by De Dreu and Gross, members are much less likely to be devoted to their groups in such circumstances *only if* the group is made up primarily of non-kin (a relatively common occurrence in the modern world).

Second, in a similar fashion, the evolutionary mismatch framework could also elucidate the underlying rationale regarding the need for, and the (limited) usefulness of, various (psychological/social and punitive) strategies indicated by the authors that were designed to induce (continued) commitment and devotion to a group. As members of many modern-day groups, as indicated previously, are posited to have a much greater tendency to experience significant fluctuations in their sense of group loyalty and commitment due to the evolutionarily novel, largely non-kin makeup of their groups, the adoption of such an assortment of different measures would understandably be required to safeguard group cohesion.

However, while such methods might be useful to a certain extent in enhancing or enforcing group loyalty, especially among deterrence-oriented groups, I dispute that they are the most effective. From an evolutionary viewpoint, tactics that attempt to address the evolutionarily mismatched nature of group makeup among many present-day groups by prominently highlighting the importance of collective action in dealing with potential threats to one's evolutionarily salient in-group members (e.g., relatives) will conceivably be much more effective (Hamilton 1964a; 1964b; Hughes 1988). For instance, a country is believed to be more likely to succeed at motivating its soldiers to stay united and be constantly prepared to deal with any future attack if it emphasizes the importance of national security in protecting the soldiers' own kin from eventual harm (as compared with the diverse strategies indicated by the authors). By the same token, measures that promote a sense of kin-like bond among members who are not otherwise biologically or affinally connected could analogously be more effective at enhancing and maintaining cohesion and loyalty (Griskevicius et al. 2012). Creating/developing the notion that a country's soldiers are all just like siblings to one another, as another corresponding example, will imaginably also be more beneficial in preserving their allegiance to the nation.

Taken together, I argue that viewing the issue of group cohesion during intergroup conflict from a novel evolutionary mismatch angle will provide some new insights into the phenomenon, which could complement the valuable information and ideas contributed by the authors and afford creative ways of improving in-group loyalty that might have important relevance for national security in the contemporary world.

# The multiple facets of psychopathy in attack and defense conflicts

Tiago O. Paiva[a,b], Rui C. Coelho[a], Rita Pasion[a], Beatriz Ribeiro[c], Pedro R. Almeida[c], Fernando Ferreira-Santos[a], João Marques-Teixeira[a] and Fernando Barbosa[a]

[a]Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences, University of Porto, 4200-135 Porto, Portugal; [b]Faculty of Medicine, University of Porto, 4200-135 Porto, Portugal and [c]School of Criminology and Interdisciplinary Research Center on Crime, Justice and Security, Faculty of Law, University of Porto, 4200-135 Porto, Portugal.
tiagopaiva@fpce.up.pt rui.cardoso.coelho@gmail.com
ritapasion@gmail.com beatriz.d.o.ribeiro@gmail.com
palmeida@direito.up.pt frsantos@fpce.up.pt jemt01@gmail.com
fbarbosa@fpce.up.pt

doi:10.1017/S0140525X19000803, e135

**Abstract**

With respect to De Dreu and Gross's article, we comment on the psychological functions for attack and defense, focusing on associations between individual differences in psychopathic personality traits and the behavioral patterns observed in attack-defense conflicts. We highlight the dimensional nature of psychopathy and formulate hypothetical associations between distinct traits, their different behavioral outcomes, and associated brain mechanisms.

When exploring the psychological functions for attack and defense, De Dreu and Gross suggest a theoretical relationship between individual psychological characteristics and behavioral profiles in attack-and-defense contexts. More specifically, the authors ultimately infer that propensities for attack and defense are intrinsically associated with the psychological characteristics of psychopathy and paranoia, respectively. This assertion seems to be based on a psychopathologically oriented conception of psychopathy that assumes psychopathy as unitary in nature. This classic view has been extensively challenged by a taxometric analysis of the distribution of self-reported psychopathic scores, suggesting psychopathy as a dimensional construct (Edens et al. 2006; Guay et al. 2007). Therefore, psychopathy is better conceptualized as a constellation of personality traits, with distinct etiological pathways (e.g., low fear and externalizing vulnerability) that may predispose individuals toward the manifestation of antisocial behaviors (Patrick et al. 2009; Woodworth & Porter 2002). It is then theoretically plausible that the distinct traits that constitute the psychopathic personality structure modulate different components of individual behavior in attack-and-defense contexts. The distinction between psychopathic traits and expected outcomes is relevant not only for formulating psychological profiles of attackers and defenders, but also for the prediction of response patterns and strategies used in asymmetric economic tasks.

The Triarchic Model of Psychopathy (Patrick 2010; Patrick & Drislane 2015; Patrick et al. 2009) defines psychopathy as a constellation of personality traits grouped into three distinct phenotypic expressions (boldness, meanness, and disinhibition) that represent biobehavioral liability factors for psychopathology (Somma et al. 2018). The expressions of boldness and meanness emerge from the same etiological dimension, which is related with the low fear models of psychopathy. According to these models, psychopathic individuals do not exhibit a fearful response to the anticipation of punishment when facing situations involving rewards and punishments that are contingent to a specific behavior (Fowles & Dindo 2006). Evidence in favor of the low fear hypothesis arises from classic psychophysiological studies showing reduced skin response to a conditioned stimulus associated with the presentation of an electric shock, along with reduced avoidance of punished responses during an avoidance learning task (e.g., see Lykken 1957). Interestingly, several brain regions implicated in aversive associative learning, and part of the integrated emotion systems, have been suggested as the core mechanism implied in psychopathy, including the central and basolateral nuclei of the amygdala, the insula, the ventromedial prefrontal cortex (VMPFC), the sensory association cortex, the posterior thalamus, and the motor cortex (Blair 2006; Blair et al. 2018). Disinhibition, on the other hand, is thought to reflect a general externalizing proneness involving lack of planfulness and foresight, poor emotional regulation, reliance on immediate gratification, and deficient behavioral restraint (Patrick et al. 2009). It is thought that externalizing vulnerability is expressed by prefrontal-based differences in inhibitory control (Fowles 2018; Patrick & Drislane 2015). In the context of attack and defense, boldness and meanness attributes are closer to an attacker profile, whereas the disinhibition phenotypic component can be expected to influence both attack and defense mechanisms.

Boldness, as a low fear disposition reflecting audacity, social dominance, and low stress reactivity, maps onto overconfidence, feelings of superiority, and reward sensitivity (Drislane & Patrick 2017). These characteristics are straightforwardly connected to the authors' descriptions of an attacker, along with the meanness that has lack of empathy as the main defining characteristic (Almeida et al. 2015). It is not obvious, however, what role disinhibition traits would have in a specific profile. On the one hand, alterations in prefrontal cortex (PFC) function may compromise strategic patterns of behavior and increase reactive attack performances through facilitation of behavioral approach and willingness for risk-taking (Blair 2007). On the other hand, when it comes to defensive profiles, disinhibition is also associated with anxiety and the related features of high sensitivity to punishment and hostility bias, predicting increased defensive behavior (Nelson & Foell 2018).

The concepts of behavioral approach and avoidance are recurring themes in the literature on psychopathy (Baskin-Sommers et al. 2010). From a behavioral and neuroeconomics standpoint, both loss aversion and reward-seeking behaviors are conceptualized in the context of motivational based behavior. At the brain level, several brain structures have been implicated in loss aversion, including the ventral striatum, the VMPFC, and the superior PFC (Tom et al. 2007). The same goes for brain activity associated with gains, which increase the activity of the dopaminergic reward circuitry that includes the dorsal and ventral striatum, VMPFC, ventrolateral PFC, and the anterior cingulate cortex (Tom et al. 2007). These regions overlap with other brain regions relevant for the etiological pathways associated with psychopathy. In this light, we may hypothesize that psychopathic traits are associated with distinct behavioral patterns in attack-and-defense contexts by modulating the utility functions associated with rewards and losses. More specifically, on the one hand, boldness and meanness are expected to decrease loss aversion, increasing the frequency of unsuccessful attack behaviors. On the other, disinhibition is expected to modulate both an increase in loss aversion (hostility bias) and reward sensitivity, increasing defensive aggression in defensive contexts (De Dreu & Kret 2016) and the disposition for attack in attack contexts, due to the lack of impulse control (Nelson & Foell 2018).

Most research on loss aversion and reward-seeking behaviors comes from game theory approaches to symmetric game interactions. In this sense, it seems fundamental to extend the relations between psychopathic traits, behavioral patterns in attack-and-defense contexts, and associated brain and etiological mechanisms in asymmetric economic games that, as suggested by the authors, are a more ecological setting for real-life conflict interactions. It is also important to rethink the proposed linear associations between psychopathy and attack proneness, considering the dimensional approaches that highlight the need for a comprehensive framework, given the heterogeneity of the psychopathic personality structure.

## Identity leadership: Managing perceptions of conflict for collective action

Philip Pärnamets[a,b], Diego A. Reinero[a], Andrea Pereira[c] and Jay J. Van Bavel[a]

[a]Department of Psychology, New York University, New York, NY 10003; [b]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, 171

77 Stockholm, Sweden and [c]Social and Organizational Psychology Unit, Institute of Psychology, Leiden University, 2300 AK Leiden, The Netherlands.
philip.parnamets@nyu.edu diego.reinero@nyu.edu
a.l.pereira@fsw.leidenuniv.nl jay.vanbavel@nyu.edu
https://philipparnamets.github.io/
https://scholar.google.com/citations?user=3Sdr9XsAAAAJ
https://scholar.google.com/citations?user=WcAughoAAAAJ
http://as.nyu.edu/content/nyu-as/as/faculty/jay-van-bavel.htm

doi:10.1017/S0140525X19000876, e136

**Abstract**

We argue that how players perceive the attack-defense game might matter far more than its actual underlying structure in determining the outcomes of intergroup conflict. Leaders can use various tactics to dynamically modify these perceptions, from collective victimization to the distortion of the perceived payoffs, with some followers being more receptive than others to such leadership tactics.

*Make them believe, that offensive operations, often times, is the surest, if not the only … means of defence.*

—George Washington (1799)

George Washington understood that social reality depends crucially on social perception. For this reason, the *perceptions* of the attack-defense dynamics might matter far more than their *actual* underlying reality in determining the outcomes of intergroup conflict. This is especially true in identity-driven conflicts where interpretations of social factors determine group cohesion and collective action, as is the case in politics. We fully agree with De Dreu and Gross (D&G) that one potent way of mustering support for attack is to create the perception that one's group is under attack. This may solve group coordination problems by bolstering in-group identity and compliance, and instituting leadership hierarchies. This shift in perception occurs, for example, when powerful groups claim they are collective victims of external forces or groups. Our commentary will highlight the role of rhetoric, persuasion, and propaganda in changing perceptions and actions during intergroup conflict.

We speculate that one characteristic of effective leaders is having an intuitive grasp of the malleability of attack-defense dynamics and the ability to manipulate their followers' perceptions (e.g., through identity leadership; Reicher et al. 2012). For example, political leaders who use collective pronouns (*we, us*) are more likely to get elected to public office (Steffens & Haslam 2013). Some leaders may go a step further and frame their side as victims – even after they win political office and control the levels of political power. For instance, a recent analysis of U.S. President Donald Trump's most viral messages on social media found that they contain collective victimization language in which he frames himself and his followers as victims (Brady et al. 2019). Thus, making one's own group feel attacked creates the perception that the attackers (e.g., with the goal of changing institutions or policy) are in fact the defenders in the conflict, and may serve as a cheap and effective means to generate a shared sense of identity and mobilize followers.

Another strategy for mobilizing followers is to alter their perception of the payoffs associated with attack versus defense. Take the situation where both groups choose to "cooperate" (CC) – attackers choose to not attack and defenders choose to not defend – which maintains the status quo. Based on D&G's model, there is some net difference, $d$, in the utilities to the groups, where the defender gains more than the attacker. However, if the attacking side can convince its members that the net difference is actually larger (formally, a transform $F(d)$ such that $F(d) > d$), this will increase the perceived value of the attackers switching to an attacking strategy on the following round. An example of this leadership strategy is commonly deployed in Western Europe among populist movements or among the alt-right in the United States to motivate conflict against out-groups. It occurs when a societal issue, such as immigration, is framed as an existential threat to the native populace. This elicits perceptions of threat and prejudice from the majority group (see Craig & Richeson 2014; Shepherd et al. 2018), and motivates action. This was heard during a Unite the Right rally in Charlottesville, Virginia, when white supremacists chanted, "You will not replace us!"

It is important to note that not all members of a group will perceive threat to the same extent (Hibbing et al. 2015). When these individual differences are unevenly distributed between two conflicting groups, one group may be more successful at mobilizing attack. Indeed, in the United States, conservatives tend to be more attuned to threats than liberals (Kanai et al. 2011; Oxley et al. 2008), and several theories argue that conservatives may place a greater premium on group loyalty and deference to authorities than liberals (Haidt 2012; Stern et al. 2014). Such individual differences might help shed light on current success of various populist movements across Western liberal democracies. Even during periods of historically low crime and relatively high prosperity, certain groups may nonetheless perceive threats that allow them to mobilize politically.

When some people perceive threats to the very existence of their group, it can elicit aggression (Wohl et al. 2012) or authoritarianism (Stenner 2005). This can occur when they perceive that rapid changes to a society, such as the racial/ethnic demographic shift within the United States, are undercutting social cohesion. Under such conditions, authoritarians feel more threat and will be more sensitive to strong identity leadership. In these cases, the normal dynamics of conflict may be reversed. Authoritarians initially may perceive themselves as defenders and later transition into the role of attackers as their cause grows stronger (Altemeyer 1988).

Accordingly, an important area for developing the theory of D&G is the dynamic nature of intergroup conflict. Using the most effective strategy requires an assessment of the current and future state of the conflict, and we expect some leaders to be better at this than others. Indeed, groups and, particularly, their leaders can use attack and defense strategies at different times or even simultaneously (as during a pre-emptive strike). The blurred nature of these strategies may also suggest that the fight-or-flight neural system, typically associated with defense, could also play a role in a successful attack – somewhat undermining the strong distinction that D&G draw between different neural systems and the roles of attacker or defender. Modeling the dynamics of repeated conflicts would also allow for group memberships to change over time, because mobilized groups might attract new members to their ranks, further changing the incentive structure of the conflict. Effective leaders will be able to exploit such changes in group identification, as when a party's membership and support ebb and flow between electoral cycles, with profound effects on the outcomes of future conflicts. Translating these dynamic processes fully into the game-theoretic

framework suggested by D&G would strengthen the appeal and generalizability of their promising account of intergroup conflict, as well as provide fertile ground for further investigation.

# Between-group attack and defence in an ecological setting: Insights from nonhuman animals

Andrew N. Radford[a], Susanne Schindler[a] and Tim W. Fawcett[b]

[a]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom and [b]Centre for Research in Animal Behaviour, University of Exeter, Devon EX4 4QG, United Kingdom.
andy.radford@bristol.ac.uk susanne.schindler@bristol.ac.uk
t.w.fawcett@exeter.ac.uk
http://www.bio.bris.ac.uk/research/behavior/Vocal_Communication/home.html
http://www.bristol.ac.uk/biology/people/susanne-schindler/overview.html
http://www.timwfawcett.com

doi:10.1017/S0140525X19000773, e137

**Abstract**

Attempts to understand the fundamental forces shaping conflict between attacking and defending groups can be hampered by a narrow focus on humans and reductionist, oversimplified modelling. Further progress depends on recognising the striking parallels in between-group conflict across the animal kingdom, harnessing the power of experimental tests in nonhuman species and modelling the eco-evolutionary feedbacks that drive attack and defence.

De Dreu and Gross (D&G) rightly address a fundamental simplification of theoretical and empirical research on between-group conflict; the prevailing assumption that competing groups are equivalent is unrealistic in many ways. By considering key differences between attackers and defenders, how those differences might arise, and the strategic, social, and psychological consequences, D&G provide an elegant initial exploration of the topic. However, we argue that two crucial extensions are needed for a full understanding of this neglected aspect of sociality and its implications for human society: a broader taxonomic perspective and a more ecologically and evolutionarily relevant modelling approach.

First, D&G focus almost exclusively on humans, ignoring similar conflicts in other social animals. Contests between rival groups are commonplace across taxa from insects to primates (Batchelor & Briffa 2011; Hardy & Briffa 2013; Kitchen & Beehner 2007; Radford 2003; Thompson et al. 2017) and nonhuman between-group conflict parallels the human scenarios discussed by D&G in many respects. For instance, fighting usually arises over a resource (e.g., food, mates, territory) already held by one group (Christensen & Radford 2018; Kitchen & Beehner 2007), and so there is an inherent role asymmetry between attackers (seeking to gain the resource) and defenders (seeking to protect that resource from usurpation). For a given pair of groups, especially in territorial species where contests between neighbours are common, which group is attacking and which is defending can change across time (birds: Radford & du Plessis 2004; mongooses: Thompson et al. 2017; primates: Wilson et al. 2012), and repeated interactions can influence subsequent behaviour even in the absence of recent conflict (Radford 2011). Moreover, collective-action problems can arise whether the resource being competed for is of direct benefit to all or only a subset of group members (Schindler & Radford 2018; Willems et al. 2015). As a final example, the home-field advantage – where defenders are disproportionately more successful than attackers – has been regularly recorded in nonhuman group-living species (primates: Crofoot et al. 2008; mongooses: Furrer et al. 2011; birds: Strong et al. 2018).

Considering only one species limits development of ideas about the causes and consequences of between-group conflict for several reasons. Conflict is a powerful selective force across taxa, but some of the conclusions drawn by D&G are likely to apply to humans only. For example, D&G emphasise the need for means to motivate and coordinate collective attacks. Whilst sanctioning systems (e.g., punishment) and rewards may occur in many species with respect to participation in between-group conflict (Arseneau-Robar et al. 2016; 2018; Radford et al. 2016), the need for additional institutional arrangements is almost certainly a uniquely human construct. Considering more species would allow the relative importance of these different mechanisms to be teased apart. Even for researchers interested only in human behaviour, our shared evolutionary history and clear parallels with other species mean there is value in taking a broader taxonomic perspective. That is particularly true because whilst D&G showcase experimental approaches that can be used with humans, these are necessarily artificial scenarios. Experimental testing of between-group conflict, including in natural conditions, is more feasible in nonhuman species: It is possible to stage contests (ants: Batchelor & Briffa 2011; fish: Bruintjes et al. 2016) and to simulate the recent (e.g., faecal presentations, mongooses: Christensen et al. 2016) and current (e.g., vocal playbacks, mongooses: Furrer et al. 2011; birds: Radford 2005; primates: Wilson et al. 2001) presence of rivals. Experimental testing alongside further theoretical development is clearly going to be key when advancing D&G's ideas concerning attack versus defence in between-group conflict.

A second limitation concerns the simple game-theory models presented by D&G, which do not capture important ecological and evolutionary forces shaping between-group conflict in real-world settings. D&G conceptualise the interaction between an attacking group and a defending group as a one-shot, simultaneous ("sealed bid") game with no variation between groups (in size, resources, fighting ability, etc.), with only two possible actions (or a one-dimensional investment in the conflict) and a pay-off matrix with fixed parameter values. This ignores the effect of future interactions on the value of the current conflict (the "shadow of the future" alluded to by D&G), the rich array of behavioural options available to group members (including retaliation after an attack), and the powerful feedbacks arising from the decisions of other groups. These are not trivial details but can fundamentally alter conflict outcomes (McNamara 2013). Work in evolutionary biology has emphasised the importance of embedding games in a wider ecological context, where the pay-offs from alternative actions emerge from the stable strategic solution to the game, rather than being specified in advance as arbitrary fixed values (Houston & McNamara 2005). D&G also ignore how contest decisions depend on uncertainty about the fighting strength of rivals, which has an important effect on

aggressive tendencies (Enquist & Leimar 1983; Johnstone 2001; McNamara & Houston 2005) and is likely to promote strategies for gathering information before launching an attack. Such considerations are particularly relevant when there are repeated interactions, which are ignored in D&G's one-shot games but are likely to be common between neighbouring groups and others in close proximity, where the threat of retaliation is high and dynamic changes in fighting strength may drive the pattern of conflict.

Whilst all models are necessarily an abstraction of reality, ignoring the wider ecological or cultural context in which between-group conflict takes place impairs our ability to derive valid, testable predictions. We urge researchers to take a broader, ecologically informed view of attack and defence, considering the rich array of examples across the animal kingdom. This is especially important when applying theoretical insights to our own species: Basing conclusions for human society on oversimplified, reductionist models is potentially dangerous, because it could misguide policy makers and politicians into accepting a limited room for manoeuvre.

**Acknowledgments.** ANR and SS were funded by a European Research Council Consolidator Grant (682253) awarded to ANR.

# Using the research on intergroup conflict in nonhuman animals to help inform patterns of human intergroup conflict

Amanda R. Ridley and Melanie O. Mirville

Centre of Evolutionary Biology, School of Animal Biology, University of Western Australia, Crawley, Perth, WA 6009, Australia.
Amanda.ridley@uwa.edu.au Melanie.mirville@research.uwa.edu.au
www.babbler-research.com

doi:10.1017/S0140525X19000827, e138

**Abstract**

There is a large body of research on conflict in nonhuman animal groups that measures the costs and benefits of intergroup conflict, and we suggest that much of this evidence is missing from De Dreu and Gross's interesting article. It is a shame this work has been missed, because it provides evidence for interesting ideas put forward in the article.

Intergroup conflict can be expected to occur whenever animals form groups, and particularly when there is variation among groups in their access to a desired resource. Given that group-living occurs in many species (Krause & Ruxton 2002), intergroup conflict is relatively common and is reviewed in Christensen and Radford (2018). Indeed, as identified by De Dreu and Gross (D&G) in their thought-provoking article, cooperation and conflict often go hand-in-hand. We applaud D&G for developing new ideas to help further our understanding of intergroup conflict in humans. We wholeheartedly agree with their criticism of existing models of human conflict, which assume symmetric values between defending and attacking groups (sect. 5, para. 1). These assumptions are not realistic, and D&G's focus on understanding asymmetries between groups is an excellent idea.

We, however, feel that the article could have been strengthened by greater consideration of the large body of research on intergroup interactions in nonhuman animals. Indeed, some of the predictions made by D&G have already been tested in animals. For example, research on primates has revealed that coordinated defense is *less* likely when the attacking individual(s) is targeting a particular individual (for example, the dominant breeding individual; Schindler & Radford 2018). Subordinate individuals may be less willing to participate in such cases of targeted attack, and opposite-sex individuals do not intervene in fights between same-sex residents and attacking individuals in a number of species (Cant et al. 2002; Ridley 2016). Therefore, the prediction that individual interests are more aligned in defending groups because they share a common fate when they lose (D&G, sect. 4.1, paras. 2 and 3) is not always supported. When a dominant individual is defeated by an attacker, for example, the opposite-sex group members can still breed with the "replacement" individual, creating a large difference between group members in the "cost" of a successful out-group attack.

The fact that individuals may vary in their contributions to intergroup conflict can also be understood by the collective action problem, where free-riders reap the benefits of successful attack or defense without incurring the costs (Nunn 2000). Recent studies have revealed considerable variation in individual contributions to intergroup conflict (Bonanni et al. 2010; Mirville et al. 2018), and these contributions are strongly related to the *individual* costs and benefits from investing in conflict. Therefore, the level of variation in benefits between individuals may determine the collective action of the group, and the benefits of investing in conflict should be considered at the individual level. This concept, which recognizes within-group conflict over defense against intruders, has recently been formally modeled by Schindler and Radford (2018). Thus, a consideration of the collective action problem would help parameterize the current models of human conflict presented by D&G.

In order to understand intergroup conflict, the motivation to invest in conflict needs to be considered. In nonhuman animals, patterns of conflict are commonly considered in relation to biological parameters. These include the sex and age of the individuals competing, and their access to resources such as food, territory, and reproduction (reviewed in Christensen & Radford 2018). Often, these biological parameters prove to be good predictors of observed patterns of conflict (Thompson et al. 2017). Although we appreciate that, at times, it is more difficult to measure some biological parameters in humans, we believe that their inclusion into human conflict research is necessary. D&G state that evidence for the causality of attacks is currently limited (sect. 5.2, para. 4). However, causality is documented in many nonhuman animal species (e.g., Thompson et al. 2017; Wilson & Wrangham 2003; among others), which can prove very informative for human studies.

Additionally, a considerable body of research on intergroup interactions in primates exists and would have provided an excellent background for some of the ideas presented in D&G's article. Chimpanzees (*Pan troglodytes*) and humans are considered to have similar social groupings and motivations for conflict (Wrangham & Glowacki 2012). Thus, the idea by D&G regarding why attack is less successful than defense would have benefitted from a consideration of the "imbalance of power" model, which has been extensively tested in chimpanzee societies (Wilson 2001; Wrangham 1999). This model assumes that one of the determinants of attack is a sufficient variation in party size

between the attacker and the defender (Wrangham 1999), and research supports this prediction, that intergroup attacks are more likely when the attacking group has numerical superiority (Wilson 2001). Although the idea of asymmetry between social groups was a central topic for D&G, evidence for the imbalance of power model was not addressed.

The prediction that intergroup conflict will affect within-group behavior in humans has been tested in a number of nonhuman animals. These studies provide considerable support for some of the predictions made by D&G. For example, there is considerable evidence that intergroup conflict increases affiliative behaviour between group members in the wood hoopoe (*Phoeniculus purpureus*) (Radford 2008), mountain gorillas (*Gorilla beringei beringei*) (Mirville 2018), and other primates (Majolo et al. 2016). This suggests that intergroup conflict may have an important influence on the dynamics of group-living behavior and would have been a useful inclusion to suggest that behavioural similarities across the animal kingdom may exist.

In summary, we found the article of D&G an extremely interesting read but felt that the article addressed primarily strategic decisions, without considering the biology behind these decisions. There is a large body of research on social (nonhuman) animal groups that has quantified the costs and benefits of intergroup conflict, and could provide support to many of the ideas put forward in the article. We encourage incorporation of some of this research into human models of intergroup conflict as a productive way to create more realistic theoretical models of asymmetric conflict.

# A note on the endogeneity of attacker and defender roles in asymmetric conflicts

Hannes Rusch[a,b] and Robert Böhm[c]

[a]Public Economics Group, School of Business and Economics, University of Marburg, 35032 Marburg, Germany; [b]School of Management, TU München, 80333 Munich, Germany and [c]School of Business and Economics, RWTH Aachen University, 52062 Aachen, Germany.
hannes.rusch@tum.de robert.boehm@rwth-aachen.de hrusch.de robertboehm.info

doi:10.1017/S0140525X19000748, e139

**Abstract**

We argue that the roles of attacker and defender in asymmetric intergroup conflict are structurally ambiguous and their perception is likely to be subjectively biased. Although this allows for endogenous selection into each role, we argue that claiming the role of the defender likely is more advantageous for conflict participants.

In their target article, De Dreu and Gross (D&G) point out important asymmetries between attack and defense that have been unduly neglected in most previous research on intergroup conflict. We generally agree with their conclusions and appreciate the comprehensive survey of theoretical and empirical results they present. However, we highlight that simplified models of asymmetric attacker-defender interactions such as those devised and discussed by D&G miss one crucial aspect of real-world intergroup conflicts: the endogenous and often subjective nature of the attacker and defender roles. The authors mention that real attack-defense relationships exhibit temporal dynamics, implying that attacker and defender roles can alternate between parties over time. Yet their analyses presume that clear identification of an attacking party and a defending party is possible. In the following, we point out three structural aspects of and two psychological phenomena observable in intergroup conflicts, all suggesting that the roles of attacker and defender are more ambiguous than claimed by D&G. We think that future work toward more comprehensive theories of the forms and functions of conflict will benefit from accounting for these ambiguities and the endogenous assumption of roles they allow.

Structurally, a first observation is that many, if not most, of the tools and skills that humans have developed for use in intergroup conflicts, that is, weapons and fighting tactics, cannot be classified as serving exclusively offensive or defensive purposes (see, e.g., Glowacki et al. 2017). Instead, they can usually be used to either end, opening enough interpretational leeway for parties to arm without sending a clear signal of offensive or defensive intentions. Second, real-world conflicts, especially those between groups, are rarely decided in a single showdown. Instead, such conflicts usually consist of a series of attacks and counterattacks, ambushes and sieges, and various other types of violent encounters that can continue for long periods. In such continuing series of skirmishes, the roles of attacker and defender are blurred and blended. Third, even in one-shot encounters in which one party clearly makes the first move, the intention of this aggressor can still be defensive: offensive action can be used to pre-emptively defend (e.g., Böhm et al. 2016; Halevy 2017; Rusch 2014a).

In addition to these structural ambiguities, at least two psychological phenomena may further bias the conflicting parties' perceptions of who is attacking and who is defending. First, individuals have reliably been found to report greater levels of fear of exploitation when interacting with groups as compared with interindividual interaction (so called "schema-based distrust"; e.g., Wildschut et al. 2003). This tendency likely amplifies defense-oriented preemptive engagement in intergroup conflicts. Second, framing the own group as the defender represents a purposeful psychological mechanism: It mobilizes greater support and decreases free-riding within the own camp *ex ante,* that is, in the light of looming conflicts (see, e.g., Abbink & Haan 2014; Doğan et al. 2018; Walker & Bailey 2013). Intriguingly, furthermore, claiming the defender role for the own group also has benefits *ex post*: It reduces the likelihood of being viewed as a depraved and/or threatening aggressor by third parties (a logic that can even result in so-called "competitive victimhood"; for a review, see Young and Sullivan 2016).

In summary, we stress that the attacker and defender roles in asymmetric intergroup conflict that D&G mostly treat as objectively and exogenously fixed are structurally more ambiguous and their perception is likely to be subjectively biased, allowing for endogenous selection into each role. Paradoxically, thus, all conflict parties can tend toward conceiving of themselves as being on the defensive side.

Investigating the psychological mechanisms underlying such self-selection into the defending role in intergroup conflicts appears as a promising direction for future research to us – not least because, as reviewed in the target article, defensive intentions

trigger a higher willingness to engage in intergroup conflicts, which thus become harder to resolve.

# The attack and defense games

Roman M. Sheremeta

Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106; Economic Science Institute, Chapman University, Orange, CA 92866.
rshereme@gmail.com https://sites.google.com/site/romansheremeta/

doi:10.1017/S0140525X19000931, e140

**Abstract**

The attack-and-defense game is a game in which an attacker (a group of attackers) has an incentive to revise the status quo and a defender (a group of defenders) wants to protect it. The asymmetry in objectives creates incompatible interests and results in a mixed-strategy Nash equilibrium. However, this equilibrium could be heavily impacted by behavioral considerations.

## *The definition of conflict*

Although conflict is a subject of study in all fields of the social sciences, the definition of *conflict* is not the same across disciplines. Economists, for example, define conflict as a situation in which competitors choose costly inputs in pursuit of private payoffs framed as wins and losses (Garfinkel & Skaperdas 2007; Kimbrough et al. 2019). Evolutionary biologists define *conflict* in terms of Darwinian fitness (Darwin 1859), where at least one of the species negatively affects the others' fitness (Rusch & Gavrilets 2019).

De Dreu and Gross (D&G) propose a definition of *conflict* as a game of attack-and-defense. Indeed, many conflicts have such structure: There is an attacker who wants to revise the status quo and a defender who wants to protect it. Also, such a definition of *conflict* conceptually fits most disciplines, by pointing out incompatibility of interests by the parties involved in the conflict.

## *Attack-and-defense games*

The main assumption behind the attack-and-defense game of D&G is that conflicting parties have asymmetric objectives. As a result of this asymmetry, the solution to such a game is a mixed-strategy Nash equilibrium.

The continuous version of the attack-and-defense game resembles an all-pay auction (contest) with externalities (Baye et al. 2012; Chowdhury & Sheremeta 2011). Also, such a game is closely connected to the attack-and-defense games of a weakest-link network of targets (Clark & Konrad 2007; Kovenock & Roberson 2018; Kovenock et al. 2019), in which the attacker's objective is to assault at least one target successfully, and the defender's objective is to defend all targets. Finally, the attack-and-defense game is closely related to games of profiling (Holt et al. 2016; Kydd 2011), in which attackers choose which demographic "type" to recruit and defenders choose which demographic types to search. All aforementioned conflict games have a structure where there is an attacker and a defender who have asymmetric objectives and whose interests are incompatible. Therefore, these games are directly linked to the attack-and-defense game of D&G.

## *Behavioral considerations*

There are many behavioral considerations that may influence the actual behavior of competitors in the game of attack-and-defense. First, if attackers are inequity averse (Fehr & Schmidt 1999), then conflicts could be less intense than predicted by the standard game theory. However, if attackers are spiteful (which is a more realistic assumption), then conflicts are more likely to escalate (Mago et al. 2016). Conflicts could also escalate if competitors, in addition to monetary utility, derive a utility from winning itself (Sheremeta 2010), or if competitors are regret averse (Filiz-Ozbay & Ozbay 2007).

Other important behavioral considerations include guilt aversion, loss aversion, overconfidence, impulsivity, and various emotional responses. Without proper game-theoretic analysis (Konrad 2009) and experimental testing (Dechenaux et al. 2015), it is not clear how these behavioral factors impact individual behavior of competitors in the game of attack-and-defense.

## *Intergroup games of attack-and-defense*

Many conflicts involve multiple attackers and defenders, resulting in an intergroup conflict game with asymmetric objectives (Chowdhury & Topolyan 2016a; Chowdhury et al. 2013). Scientists from different fields have been fascinated by such games not only because of their prevalence in real life, but also because intergroup conflicts have a number of interesting features with non-trivial trade-offs. For example, attackers may have an incentive to cooperate with each other by expending effort in order to carry out a successful attack; however, because effort is costly, each attacker also has an incentive to abstain from expending any effort and instead free-ride on the efforts of others. Also, because attackers and defenders have asymmetric objectives, the problem of free-riding can be different for the group of attackers and the group of defenders. D&G point out that defender groups share a common fate when they lose, and so their individual interests are more aligned than those of attacker groups. Consequently, this asymmetry results in stronger incentives to free-ride among attackers than among defenders. Another interesting asymmetry pointed out by D&G is that in-group identification could be stronger among defenders than attackers. These are interesting hypotheses that are worth further investigation.

## *Mechanisms of intragroup cooperation*

The asymmetry in free-riding incentives between attackers and defenders creates a greater need for attackers to use various cooperation mechanisms in solving the free-riding problem. Such mechanisms could involve negative reinforcement, such as shaming, sanctioning, punishment, and ostracism, as well as positive reinforcement, such as communication, leadership, and feedback. Although most of these mechanisms have been studied in the context of intergroup conflicts (Sheremeta 2018), they have not been studied in the context of intergroup attack-and-defense games. This is an interesting avenue for future research.

# Advantaged- and disadvantaged-group members have motivations similar to those of defenders and attackers, but their psychological characteristics are fundamentally different

Nurit Shnabel[a] and Julia Becker[b]

[a]The School of Psychological Sciences, Tel-Aviv University, Israel 69978 and [b]Institute of Psychology, University of Osnabrueck, D-49074 Osnabrueck, Germany.
shnabeln@tauex.tau.ac.il julia.becker@uni-osnabrueck.de
https://socsci3.tau.ac.il/nurit-shnebel/
https://www.psycho.uni-osnabrueck.de/en/subfields/social_psychology.html

doi:10.1017/S0140525X19000736, e141

**Abstract**

Modern societies are characterized by group-based hierarchies. Similar to attackers, disadvantaged-group members wish to change the status quo; like defenders, advantaged-group members wish to protect it. However, the psychological arrays that are typical of disadvantaged- and advantaged-group members are opposite to those of attackers and defenders – suggesting that the Attacker-Defender Game does not capture the dynamics between advantaged and disadvantaged groups.

By putting forward the Attacker-Defender Game (AD-G), the target articles makes an important contribution to behavioral game theory, which so far has neglected asymmetric games. While we applaud this contribution, the purpose of our commentary is to point to one type of asymmetric conflict that is central to modern societies yet is not captured by the model.

The AD-G models the distinct roles of *attackers*, who are motivated to increase their gains by changing the status quo, and *defenders*, who are motivated to defend against losses by protecting the status quo. The different psychological arrays that characterize attackers and defenders are said to be linked to evolved psychological mechanisms that are also evident in nonhuman animals (i.e., predator and prey). However, a major characteristic of modern societies, which is absent from societies of nonhuman animals or hunters-gatherers, is that they produce stable economic surplus, leading to the creation of group-based hierarchies (based on ethnicity, nationality, class, etc.) (Sidanius & Pratto 1999).

In these societies, disadvantaged-group members are typically more strongly motivated to change the status quo toward greater equality, whereas advantaged-group members are typically more strongly motivated to protect the status quo that privileges them (Saguy et al. 2008). Thus, the motivations of disadvantaged-group members and attackers (i.e., change the status quo) and advantaged-group members and defenders (i.e., protect the status quo) are basically similar. Whereas the authors acknowledge that "asymmetry in power … can dramatically change the motivation to attack or to defend" (De Dreu & Gross, sect. 2.3, para. 2), we argue that power asymmetry creates *substantially different* psychological arrays. Specifically, we suggest that the psychological arrays that are typical of members of disadvantaged and advantaged groups are quite opposite to those of attackers and defenders (respectively). We now turn to illustrate these opposing patterns in six examples.

First, whereas in-group identification is stronger among defender than attacker groups, disadvantaged-group members, who often confront prejudice and discrimination due to their group affiliation, are typically more strongly identified with their groups than advantaged-group members (Simon & Brown 1987). In fact, because they rarely confront prejudice and discrimination, advantaged-group members are sometimes oblivious to their group identity (e.g., whites who profess to be color-blind; Helms 1990).

Second, whereas feeling superior is functional for attackers more than for defenders, superiority beliefs (e.g., in the form of racial or religious supremacy ideologies) are more common among members of advantaged versus disadvantaged groups (a pattern termed *ideological asymmetry*; Sidanius & Pratto 1999). In fact, disadvantaged-group members sometimes internalize the stigma on their group (e.g., sexual minorities who accept homophobic attitudes as legitimate; Herek & McLemore 2013).

Third, attackers, but not defenders, are likely to be overconfident (i.e., overestimate their relative strength). Yet, group efficacy (the belief that one's group has control over and can change society; Mummendey et al. 1999) is typically stronger among advantaged groups compared with disadvantaged groups. To illustrate, individuals with lower (vs. higher) subjective social class showed cultural practices (including interests in education, arts, newspapers, TV, and shopping) that relate to recognizing the in-group's low group efficacy and, in turn, remained politically inactive when faced with an in-group–related social disadvantage (Becker et al. 2017).

Fourth, members of defender groups are more likely to cooperate with each other than members of attacker groups, who are more likely to show self-serving behaviors (e.g., free-riding). Yet, research on social class shows that members of disadvantaged groups (i.e., lower class individuals), who must rely on mutual aid, show more communal, prosocial behavior toward others than members of advantaged groups (higher class individuals), who enjoy greater independence and show more agentic, self-serving behavior (Kraus et al. 2012).

Fifth, attack is associated with activation of neural circuitries involved in the processing of rewards, whereas defense is associated with neural circuitries involved in threat detection and risk avoidance. By contrast, because elevated power activates approach-related tendencies, whereas reduced power activates inhibition-related tendencies (Keltner et al. 2003), advantaged-group members are more attentive to rewards, whereas disadvantaged-group members are more attentive to punishments and threats in their environment (Rucker et al. 2018).

Sixth, because attack means that targets may be harmed, subordinated, and exploited, attackers are more likely to experience guilt than defenders. However, especially when reminded of social inequality, advantaged-group members are more likely to experience group-based guilt than disadvantaged-group members. For example, when exposed to information about gender inequality, men expressed greater moral shame and wished that their in-group would behave more morally than women (Hässler et al. 2018).

We conclude that although members of attacker and disadvantaged groups, on the one hand, and members of defender and advantaged groups, on the other hand, share the same basic motivation pertaining to the status quo, the AD-G should not be viewed as a representation of the dynamics between advantaged

and disadvantaged groups. *Why is it important to explicitly note this?* The representations, framings, and language used by social scientists can inform and influence the public discourse about the topics that they study. Hence, our purpose in this commentary is to highlight that disadvantaged and advantaged groups should not be equated with attackers and defenders, and the relations between disadvantaged and advantaged groups should not be described as "a clash between one side seeking change and … victory, and the other side … protecting against loss and defeat" (De Dreu & Gross, abstract). Rather, as opposed to the AD-G's utility matrix, a disadvantaged group's success in changing the status quo does not necessarily involve a loss to the advantaged group. For example, from a broad historical perspective, white Americans did not "lose" because of the abolishment of slavery and racial segregation, and men did not "lose" because of the suffrage movement's success in achieving voting rights for women. Concurring with Paulo Freire's (1970) notion that "the great humanistic task of the oppressed is to liberate themselves and their oppressors," we argue that changing the status quo toward greater equality may benefit both the advantaged and the disadvantaged – by allowing individuals to enjoy secure and positive social identities (Nadler & Shnabel 2015) and live to their full potential to the benefit of society as a whole.

# Levels of analysis and problems of evidential support in the study of asymmetric conflict

Dragos Simandan

Geography Department, Brock University, St. Catharines, ON, Canada, L2S 3A1.
simandan@brocku.ca
https://brocku.ca/social-sciences/geography/faculty-and-staff/dragos-simandan/

doi:10.1017/S0140525X19000682, e142

**Abstract**

The contribution by De Dreu and Gross oversimplifies the complexity of the topic. I provide counterarguments that undermine the two sweeping contentions on which the article's argument depends, and I argue that asymmetric conflict is best understood at the finer-grained level of studying the sequences of strikes and counterstrikes that the rival actors have in store for one another.

De Dreu and Gross provide a welcome contribution to the theory of conflict by questioning the assumption of symmetry pervading much of the existing game theoretical literature. They bring to the foreground of current debates the significance of distinguishing between attack and defense and demonstrate this point through a wide-ranging review of neurobiological, psychological, and cultural mechanisms associated with this distinction. These merits notwithstanding, their contribution oversimplifies the complexity of the topic in several distinct ways that, taken together, cast doubt in the theoretical and practical insights of their proposal.

In this commentary, I show that counterexamples can be adduced to undermine two key sweeping contentions on which the target article's argument depends. To begin with, the claim that "group-level defense creates a common fate for defenders that is absent in attackers" (sect. 4.5, para. 1) is a generalization that ignores the obvious fact that "the defenders" are never a homogeneous group in all respects. There are multiple axes of social difference, including, but not limited to, race, ethnicity, religion, gender, sexual orientation, age, (dis)ability, and social class (Simandan 2019d). These axes of social difference induce profound heterogeneities in the specific "fates" that the various subgroups constituting the higher-order grouping of the "defenders" will face. To illustrate, given that Nazi ideology specifically targeted the elimination of Jews; Roma; disabled people; and lesbian, gay, bisexual, and transgender (commonly referred to as *LGBT*) minorities, the occupation by Nazi Germany of large swaths of Europe during World War II resulted in very different outcomes for these ideologically targeted minorities compared with the less "problematic" ethnic majorities of the occupied territories (Childers 2018). To give another illustration representative of the ethnic fragmentation associated with the political geography of the nation-state, Transylvania has belonged to Romania for the last hundred years, but, before that, it was a part of the Habsburg and then Austrian-Hungarian empires (Treptow 1997). Even though Romanians constitute the ethnic majority, the province has a substantial Hungarian minority. Given what we know about homophily and ethnocentrism (Bizumic 2019; Currarini et al. 2009; Jones 2018; Salter 2008; Stavenhagen 2016), it strains credulity to suggest that if Hungary were to invade Transylvania, the fate of the occupied would be a "common fate," regardless of whether they are the Romanian majority or the Hungarian minority.

The second and equally problematic sweeping contention on which the target article's argument depends is the claim that "the negative consequences of failed defense are stronger and more extreme than the consequences of failed attack" (sect. 5.1, para. 4). On the one hand, this second claim presupposes and therefore reproduces the questionable assumptions about the alleged common fate of "defenders" of the first claim. On the other hand, the historical and military records suggest that the relative severity of failed defense *versus* failed attack depends on the contextual specificities of the conflict under investigation. To illustrate, a viable competitive strategy is for a party to act weak and/or oblivious so as to bait its rival into a rushed, overconfident attack (Freedman 2015). Because the attack wasn't surprising at all, the defenders can mount a counterattack that can often be devastating for the original attackers on two grounds: firstly, the intelligence and foresight of the defenders can give them time to orchestrate a well-thought-out counterattack; secondly, because the initial attack was induced by the defenders' tactic of appearing weak and/or oblivious, the powerful counterattack is especially likely to take them by surprise and to find them unprepared and vulnerable (see also Simandan 2010; 2018b; 2019a).

I also argue that the manner in which the authors model attack and defense as games of strategy is misleading to the extent that it does not take into account the optimal level of analysis at which such modelling should take place. More specifically, asymmetric conflict is best understood at the finer-grained level of studying the sequences of strikes and counterstrikes that the rival actors have in store for one another (Simandan 2018a; 2019b; 2019c). In other words, it is less productive for the study of conflict to think in terms of attackers *versus* defenders than to think in terms of the specific chains of moves and countermoves that, taken together, constitute the higher-order "conflict." De Dreu and Gross mention only in passing this micro-level of analysis (sect. 3.4), and this analytical oversight severely circumscribes the

range of insight that their current framework can offer. This problem should be remedied in their future work by more carefully articulating the study of conflict at finer-grained levels of analysis. As a constructive suggestion of how this task could be carried out, I end this commentary by briefly delineating four complementary criteria for classifying move/countermove pairs (for details, see Simandan 2018a; 2019b; 2019c). The first criterion is *intentionality,* and its application allows us to appreciate the fact that counterforce creation does not require conscious decision-making, and that, therefore, we can usefully distinguish intended countermoves from unintended counterforces. A second criterion that carries significant analytical traction in characterizing move-countermove dyads is the degree of *similarity* between the substantive, intrinsic features of the initial move and the properties of the subsequent response. Its application yields two broad categories: similar (or symmetric) countermoves, which describe responses that are of the same kind as the triggering move, and dissimilar (or asymmetric) countermoves, which refer to reactions that are substantively different from the initial trigger. The third criterion by which move-countermove pairs can be usefully classified is the degree of *concentration of human agency* involved. One can thereby distinguish between individual countermoves and collective or diffuse responses. This distinction is significant for theoretical and methodological reasons in both the social sciences and historiography. Finally, the fourth criterion is *the time elapsed* between the initial move and the countermove. The distinction of immediate countermoves from delayed countermoves (1) brings out the complication that even immediate responses cannot happen instantaneously, (2) prompts the further classification of delays themselves into unavoidable and deliberate delays, and (3) opens questions about the advantages of making use of deliberate delays when crafting one's reaction to a competitive challenge.

# Using political sanctions to discourage intergroup attacks: Social identity and authority legitimacy

Karolina Urbanska[a] and Sam Pehrson[b]

[a]Department of Psychology, University of Sheffield, Sheffield S1 2LT, United Kingdom and [b]School of Psychology and Neuroscience, University of St. Andrews, St. Andrews KY16 9JP, United Kingdom.
karolina.urbanska@uca.fr sdp21@st-andrews.ac.uk
https://kurbanska.wordpress.com/
https://risweb.st-andrews.ac.uk/portal/en/persons/samuel-david-pehrson%283f30535e-da67-4f49-be2a-636671eff92a%29.html

doi:10.1017/S0140525X1900075X, e143

**Abstract**

De Dreu and Gross offer novel solutions to discouraging attackers via political sanctions. We offer insights from social psychological and criminological research on when such sanctions would work and when they could backfire. We argue that the influence of such sanctioning ultimately rests upon the extent to which such authorities can claim to represent the society that they serve.

Authorities have a distinct position in our societies in that a mere appointment to such duty warrants a form of power and influence over society's behaviours. People usually obey authorities if they perceive them as legitimate (Tyler 2006), and policies and reforms introduced by political authorities have a direct influence on social norms (Guimond et al. 2013; Tankard & Paluck 2016). In their article, De Dreu and Gross (D&G) suggest that authorities are among solutions to prevent grave intergroup conflict by introducing political sanctions, discouraging potential attackers. In this commentary, we elaborate on this argument by considering when political sanctioning can be effective in preventing conflict and when it could backfire, provoking reactive attacks further. We argue that the extent to which authority is perceived to represent people whom they are meant to serve is key in the extent to which political sanctions are effective.

On one level, the effect of political sanctioning on behaviour seems straightforward: sanctions discourage behaviours by altering the cost-benefit considerations such that unwanted behaviours become more costly to the actor. From this point of view, it does not matter what kind of relationship exists between the authority and potential attackers, as long as the latter is afraid of losing resources valuable to their group if they disobey the former. These ideas go back to the early theories on equity (Adams 1966; Messick & Cook 1983), whereby people are considered to be rationally weigh potential gains to one's own investments in an exchange. Judgements of authority resource distributions, in this perspective, are underpinned by the crude input-output exchanges. As D&G acknowledge, attacks tend to be less coordinated and, in turn, riskier; having another danger relating to the subsequent sanction could indeed decrease the likelihood of an attack. Thus, if instrumental motives were the sole basis for complying with authorities, then people would comply when the promise of rewards and threat of sanctions are such that compliance maximises benefits (or minimises loss).

But are authorities simply sources of rewards and punishments? Tyler and colleagues have long argued that people obey authorities that treat them fairly for relational rather than instrumental reasons; such fair and respectful treatment tells them something about their social standing in the society (Tyler & Lind 1992). Similarly, the threat of sanctions alters not only the cost-benefit implications of an action, but also our understanding of the relevant social relationships. For example, through being forced to follow a course of action against our will, we may come to see our relationship with an authority as unequal and conflictual, which, in turn, will make it difficult for the authority to subsequently appeal to a sense of duty to obey (Turner 2005). Moreover, if attack behaviour serves to enact a group identity and understanding of the world, then sanctions will not necessarily counteract the motivation for it in the way that a simple cost-benefit account would suggest. For example, in the case of a terrorist group whose very *raison d'être* is rooted in the perceived illegitimacy of the status quo, sanctioning by the authorities would be entirely in line with the worldview from which their violence derives its meaning. Indeed, as Turner also points out, to be punished by an illegitimate authority could become a badge of honour, just as a promise of reward can be resented as an attempt to control through bribery. The point is that whether a reward or sanction serves to encourage or discourage a behaviour has as much to do with the social relationship within which it is administered as with any kind of inherent value. Moreover, to the extent that sanctions are felt by a wider group of people than those already engaged in conflict, there is the potential for escalation

as is seen most clearly in studies of "public order" policing (e.g., Reicher 1996).

If the authority imposing a political sanction is viewed as being aligned or partial to an outgroup, then attackers are unlikely to identify with these decision-makers and will claim that this institution does not serve their interests (Pehrson et al. 2017; Radburn et al. 2016). Despite the risk of a sanction, lack of identification is a basis for rejecting unfavourable decisions (Huo et al. 1996). On the other hand, if the institutions are viewed by potential attackers as representative of a wider ingroup, then the picture is more optimistic. Because identification with an authority forms the basis for their legitimacy, it is more likely that even sanctions could be accepted as long as the authority represents a wider ingroup and is therefore legitimate. Thus, in settings where an authority is genuinely accepted as an "honest broker" in managing competing intergroup interests, and thus able to secure the interests and loyalty of multiple parties by keeping the peace, then it does indeed have an important role to place in minimizing attack whether between individuals or between groups. On the other hand, when the fate of the group is on the line and ingroup authorities act in a way that is seen as partisan, sanctioning can have devastating effects. Processes underpinning identification with authorities are dynamic, and thus, even where an authority is initially viewed as part of an ingroup, sanctioning may well disrupt this (Radburn & Stott 2018).

Taking these insights into account, we extend the implications of D&G by adding that policymakers wishing to discuss the effectiveness of political sanctioning in discouraging attacks should carefully assess the source or the institution of the sanctioning (whether it is understood by the targets of sanction to be representative of ingroup, outgroup, or superordinate interests) and how the sanctioning itself would affect this. Ultimately, efforts to foster high levels of identification with the political authorities should be deployed to ensure that political sanctioning does not fail and even intensify the desire to attack.

# Do people always invest less in attack than defense? Possible qualifying factors

Ori Weisel

Coller School of Management, Tel Aviv University, Tel Aviv, Israel, 6997801.
oriw@tauex.tau.ac.il https://en-coller.tau.ac.il/profile/oriw

doi:10.1017/S0140525X19000955, e144

**Abstract**

In many conflict situations, defense is easier to mobilize than attack. However, a number of factors, namely, the initial endowments available to each side, the stakes of the conflict, the respective costs of defense and attack, and the way that conflict is framed and perceived, may make attacking more attractive than defending.

De Dreu and Gross convincingly argue that symmetric conflict is the exception rather than the rule; conflict often involves attackers and defenders, rather than two equally placed parties. One of the central results of their investigation is that "people invest less in attack than defense and attack often fails" (abstract). Here, I point out a number of idiosyncratic features of the Attacker-Defender Game (AD-G), introduced and used by the authors to compare the respective propensities for attack and for defense, which may limit the generalizability of this result.

The authors introduce the AD-G as a model of asymmetric conflict. Each player – the attacker and the defender – decides whether or how much to invest in attack/defense. A successful attack occurs when the attacker invests more than the defender. The AD-G has two essential features: (1) Mutual cooperation is the most attractive option for defenders, but attackers are better off following a successful attack. (2) Defenders benefit from matching attackers' investments, whereas attackers benefit from mismatching defenders' investments.

First, I draw attention to the initial endowments, or power, available to attackers and defenders. In reality, these can greatly differ, for example, in conflict between terrorists and a state (low-power attacker vs. high-power defender), or between a multi-branch chain and a local shop the chain is trying to put out of business (high-power attacker vs. low-power defender). Although De Dreu and Gross acknowledge that "Asymmetry in power can be modeled by inequality in resource endowments in the AD-G contest game and can dramatically change the motivation to attack or to defend" (sect. 2.3, para. 2), the experiments they report on, in which they found that defenders invest more than attackers, feature equal endowments for attackers and defenders. Would results be different if endowments were not equal? This is more than just an interesting follow-up question; a positive answer implies that a main thesis of the target article is qualified by the allocation of endowments.

A related issue is that defenders in the AD-G are left with nothing following a successful attack. This design choice reflects the Dawkins and Krebs quote that opens section 3: "A rabbit runs faster than the fox, because the rabbit is running for his life while the fox is only running for his dinner." But conflict isn't always about a defender running for his or her life. The stakes could be much smaller; for example, a revisionist state may seek to capture just a small part of a neighboring state's territory. It is reasonable to assume that defenders' willingness to invest in defense will increase in the ratio between what they stand to lose (the stakes) and what they possess to begin with. For the rabbit escaping a fox, the ratio is 1; for the neighboring state, it is much lower.

The AD-G, designed to model asymmetric conflict, retains a degree of symmetry with respect to the costs of attack and defense. To survive an attack, defenders need to invest at least as much as attackers. However, in reality the cost of defense could be either lower or higher than that of attack (e.g., when defending a fortified city or scattered villages, respectively). To provide initial insights on how respective costs of attack and defense (as well as stakes and initial endowments) affect the propensity to invest in attack and defense, I analyze a generalized version of the AD-G – the Generalized Attacker-Defender Game (GAD-G; Fig. 1) – which allows for freedom in a number of important parameters: (1) Defenders and attackers can have different endowments ($e_{\text{def}}$ and $e_{\text{att}}$); (2) the stake(s) does not necessarily equal the defender's entire endowment ($s \leq e_{\text{def}}$); and (3) the cost of attack ($c_{\text{att}}$) is independent from the cost of defense ($c_{\text{def}}$).

Assuming $c_{\text{att}} < s$ and $c_{\text{def}} < s$, the GAD-G retains the essential features of the AD-G (see para. 2). Like the AD-G, there is no Nash equilibrium in pure strategies. In a mixed strategy

(a)

| | | Attacker | |
|---|---|---|---|
| | | Cooperate | Attack |
| Defender | Cooperate | $e_{def}$ \| $e_{att}$ | $e_{def}-s$ \| $e_{att}-c_{att}+s$ |
| | Defend | $e_{def}-c_{def}$ \| $e_{att}$ | $e_{def}-c_{def}$ \| $e_{att}-c_{att}$ |

(b)

| | | Attacker | |
|---|---|---|---|
| | | Cooperate | Attack |
| Defender | Cooperate | 2 \| 1 | 0 \| 2 |
| | Defend | 1 \| 1 | 1 \| 0 |

**Figure 1.** (Weisel) (a) Generalized Attacker-Defender Game (GAD-G). $e_{def}$ = defender endowment; $e_{att}$ = attacker endowment; $s$ = stakes; $c_{def}$ = defense cost; $c_{att}$ = attack cost. (b) When $e_{def} = s = 2$, $e_{att} = 1$, and $c_{att} = c_{def} = 1$, the GAD-G is identical to the ordinal version of the AD-G presented by De Dreu and Gross (sect. 2.1, para. 1).

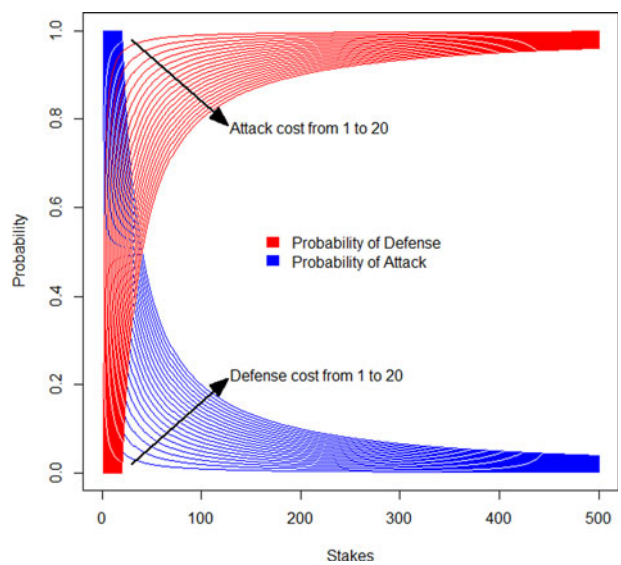


**Figure 2.** (Weisel) Mixed strategy equilibrium probabilities of attack and defense in the GAD-G. Each blue (*red*) curve indicates, for a given cost of defense (attack), the probability (in a mixed strategy equilibrium) that the attacker (defender) will choose to attack (defend), as a function of the stakes. Note that the probability of attack depends on the cost of defense (and on the stakes), and vice versa.

equilibrium, the respective probabilities of attack and defense are $c_{def}/s$ and $s - c_{att}/s$. These probabilities, for different stakes and different costs, are plotted in Figure 2 (note that the endowments $e_{def}$ and $e_{att}$ do not affect the mixed strategy equilibrium). It is evident that (i) as stakes grow larger, defense is more likely and attack is less likely; and (ii) attack (defense) is more (less) likely as the cost of defense (attack) increases. Crucially for the discussion here, when the stakes are low, the probability of attack in equilibrium is higher than the probability of defense, especially when defense and attack are expensive. If this intuition bares out in actual behavior, the particular stakes and the costs of attack and defense further qualify the result that defense is more likely than attack.

My last point has to do with the way (asymmetric) intergroup conflict is framed and perceived. Weisel and Zultan (2016) examine an asymmetric version of the Intergroup Prisoner's Dilemma (IPD; Bornstein 2003). The game models conflict between attacker and victim (i.e., defender) groups. Members of the attacker group choose between keeping resources to themselves and contributing to their group at a personal cost. Contributions benefit the in-group and simultaneously harm the victim group. Members of the victim group face a similar choice, but their contributions do not affect the attacker group. Weisel and Zultan manipulate the way conflict is framed and perceived. In the *Comparison* frame, payoffs were expressed as a function of the difference between contributions in each group. In the *Individual Harm* frame, payoffs were expressed as a function of individual choices. The two frames are equivalent in terms of actual payoffs. The Comparison frame is similar to how De Dreu et al. (2016a) describe the Intergroup AD-Contest Game, which they use to test asymmetric intergroup conflict. The results are similar as well; in both cases, defenders contribute more than attackers. Strikingly, Weisel and Zultan obtain the opposite pattern under the Individual Harm frame, where attackers invested more than defenders, suggesting that the way conflict is framed and perceived may crucially undermine the increased propensity for defense over attack.

# Authors' Response

## Asymmetric conflict: Structures, strategies, and settlement

Carsten K. W. De Dreu[a,b] and Jörg Gross[a]

[a]Institute of Psychology, Leiden University, 2300 RB Leiden, The Netherlands and [b]Center for Research in Experimental Economics and Political Decision Making (CREED), University of Amsterdam, 1018 WB Amsterdam, The Netherlands.
c.k.w.de.dreu@fsw.leidenuniv.nl mail@joerg-gross.net
https://www.universiteitleiden.nl/en/staffmembers/carsten-de-dreu
http://www.joerg-gross.net

doi:10.1017/S0140525X1900116X, e145

**Abstract**

Our target article modeled conflict within and between groups as an asymmetric game of strategy and developed a framework

to explain the evolved neurobiological, psychological, and sociocultural mechanisms underlying attack and defense. Twenty-seven commentaries add insights from diverse disciplines, such as animal biology, evolutionary game theory, human neuroscience, psychology, anthropology, and political science, that collectively extend and supplement this model in three ways. Here we draw attention to the superordinate structure of attack and defense, and its subordinate means to meet the end of status quo maintenance versus change, and we discuss (1) how variations in conflict structure and power disparities between antagonists can impact strategy selection and behavior during attack and defense; (2) how the positions of attack and defense emerge endogenously and are subject to rhetoric and propaganda; and (3) how psychological and economic interventions can transform attacker-defender conflicts into coordination games that allow mutual gains and dispute resolution.

## R1. Introduction

A substantial majority of past and present conflicts are about something owned by one and desired by another. These are the territorial struggles among nation states, the tribal raids for cattle, the neighborhood conflicts about parking spaces and barking dogs, and the board room battles for status. As such, human conflicts share many of the structural properties seen in conflicts among nonhuman animals, including the border patrols by groups of chimpanzees, shouting games between groups of territorial birds, between the lion and the wildebeest, even between viruses and their host's immune system. Yet when it comes to human conflict, theory and research heavily focused on symmetric conflicts and largely ignored the asymmetric nature of those conflicts in which one party seeks change and revision and the other party seeks to maintain the status quo.

Our target article, therefore, examined the possible structural, neuropsychological, and sociocultural aspects of attacker-defender conflicts within and between groups of people. Twenty-seven commentaries from evolutionary and animal biology, human neuropsychology, anthropology, experimental economics, psychology, and the political sciences largely resonated with our perspective and add important new insights and ideas (see Table R1). Alone and in combination, the commentaries complement and extend our approach, and they offer a range of new hypotheses and possible strategies for conflict resolution and peace settlement. We discuss these insights and extensions in relation to (1) the structure and strategy of asymmetric conflict (sect. R2); (2) the emergence and enactment of attacker and defender positions, with implications for group identification and leadership (sect. R3); and (3) possible interventions that transform attacker-defender conflicts into mutual gains bargaining amenable for dispute resolution (sect. R4). Section R5 concludes.

## R2. The structure and strategy of attacker-defender conflict

We modeled attacker-defender conflicts as an asymmetric game in which one party (attacker) competes to increase its gain and another party (defender) competes to protect against loss (**Chowdhury**; **Sheretema**; **Weisel**). Modeling conflict as an asymmetric game of strategy is neither believed nor intended to innovate game theory. It does, however, innovate conflict theory, generating novel hypotheses about the neural, psychological, and sociocultural mechanisms that operate during conflict, leading to better prediction of action tendencies and strategic maneuvering during conflict, and new ways of dispute resolution and conflict settlement.

Before moving to specific insights and extensions, two issues need to be clarified. First, we neither dismiss nor intended to devalue extant work on symmetric conflict (**Huffmeier & Mazei**). Yet, while we believe this earlier work can be insightful and of great help, we have argued that much of the work on symmetric conflicts cannot be extrapolated to conflicts between those who defend the status quo and those who seek to change it (**Mifune & Simunovic**; **Weisel**). Second, an asymmetric conflict model, first and foremost, helps identify the *superordinate goals* that antagonists have, with some wanting to keep what they have (viz., the status quo) and some wanting to take away what others have (viz., changing the status quo in its favor; also see Weisel). To achieve its superordinate goal of maintaining versus changing the status quo, antagonists have a range of *strategies and tactics* available. To defend the status quo, individuals and their groups may resort to pre-emptive strikes and pro-actively attack their revisionist aggressors. Such offensive actions serve as a *means* to protect and defend the status quo. Likewise, attackers may vigilantly protect their resources for attack. Such

**Table R1.** Summary of main topics and issues raised across all commentaries

| Target Article[a] | Topics Raised[b,c] | Commentary |
|---|---|---|
| Structure (1,2,4) | Extended forms and basic features of attacker-defender conflicts (R2) | Chowdhury; Krawcyck; Mifune & Simunovic; Radford et al.; Sheretema; Weisel |
| | Dependence and coercive power (R2) | Andrews et al.; Buckner & Glowacki; Fog; Halevy; Huffmeier & Mazei; Radford et al.; Simandan; Shnabel & Becker; Weisel |
| | Strategy selection; tactical maneuvering (R2) | Buckner & Glowacki; Lopez; Radford et al.; Ridley & Mirville; Weisel |
| Strategies and processes (3,4) | Neuropsychological mechanisms and personality (R2, R3) | Hurlemann & Marsh; McLoughlin & Corriveau; McNaugthon & Corr; Paiva et al. |
| | Role endogeneity; framing (R3) | Andrews et al.; Becker & Dubbs; Hafer; Lopez; Rusch & Böhm |
| | (Regulating) Group identity (R2, R3) | Fog; Katna & Cheon; Marie; O; Pärnamets et al. |
| Settlement (5) | Negotiation (R4) | Halevy; Huffmeier & Mazei; Urbanska & Pherson |
| | Emotion regulation (R4) | Cernadas Curotto et al.; Sheretema; Urbanska & Pherson |

[a]Main sections in the target article.
[b]Listed here are only topics that emerged across several commentaries.
[c]Numbers preceded by R refer to the relevant section in the response article.

protective measures serve as a *means* to change the status quo in one's favor. Thus, in theory, the very same action – a preemptive strike, staying on guard, or creating political alliances – can serve the distinctly different goals of protection and defense, or seeking to change the status quo.

In the interest of parsimony, our basic asymmetric attack-defense model largely ignored structural features of the conflict that can be of great influence. One such feature is the presence or absence of an explicit reference point that defines the status quo; our binary AD-G lacks such an explicit reference point, although it is clearly defined in the AD-G contest version (also see **Chowdhury**). Our commentaries highlight several other *structural features*, most notably the probabilistic nature of conflict outcomes, the (un)availability of disengagement, and differences in coercive power between the attacker and the defender. We address these first and then discuss the *means* available for tactical maneuvering and strategy selection during attack and defense, including the matching-mismatching of strategies, and the timing and sequencing of moves and countermoves.

### R2.1. Deterministic versus probabilistic conflict outcomes

Similar to related attacker-defender conflict games, we modeled asymmetric conflicts on the basis of the assumption that conflict outcomes are deterministic, defined by the strength of attack relative to defense (viz., all-pay auctions). Sometimes, however, conflict outcomes are probabilistic. Even when attack is more (versus less) powerful, defenders still survive (or are, nevertheless, defeated) (**Chowdhury).** Such "noise" can have many causes, including equipment failures and unforeseen environmental incidences. We share Chowdhury's intuition that (groups of) individuals may strategize and invest in conflict differently when outcomes are probabilistic rather than deterministic. **Buckner & Glowacki**'s analysis of raiding parties even suggests that environmental incidences, like anticipated rainfall or darkness, are sometimes factored in when designing attack strategies and that doing so can substantially increase the attacker's success-rate. The AD-G can be modified to capture these intuitions by modeling the outcome of the contest as a lottery (see, e.g., Lacomba et al. 2014). In this case, investments of Party A ($c_A$) increase the relative chance to succeed against Party B: $p_A = \lambda c_A/(\lambda c_A + c_B)$, and vice versa, $p_B = 1 - p_A$. The lambda parameter captures a (dis)advantage of the invested resources of one party over another (e.g., rainfall being more advantageous for attackers), which is equivalent to an asymmetry in available resources across parties, creating a paradox of power (Hirshleifer 1991). Risk-tolerance and loss aversion (Chowdhury), along with related constructs such as overconfidence and vigilance (see sect. 3 of the target article), are likely candidates that influence the behavior when conflict outcomes are probabilistic rather than deterministic, opening up interesting avenues for future research in asymmetric conflicts.

### R2.2. Power to disengage and to coerce

Our target article focused on conflicts without options for so-called disengagement. In the AD-G, attackers can choose to attack more or less forcefully, and defenders can choose to invest more or less in defense. In contests (e.g., the AD-G with continuous action space), such conflict expenditures model the effort that antagonists invest in their goal pursuit (i.e., victory or survival). Theoretically, such conflict expenditures can reflect the number of troops being mobilized, the mounting of defensive structures, or the metabolic energy spent on, for example, running away. Nonetheless, commentators correctly note that antagonists oftentimes have or create additional options, including those for disengagement. Such disengagement options have been built into games of strategy. A good example is the PD-Alt (**Huffmeier & Mazei**; Miller & Holmes 1975) in which antagonists can choose the "withdrawal" option that secures better outcomes than unilateral cooperation but worse outcomes than unilateral competition. Antagonists opting for such withdrawal thus reduce interdependency (Bacharach & Lawler 1981; Giebels et al. 2000), protecting against the risk of being exploited but also foregoing the benefits of mutual cooperation or exploitation (Gross & De Dreu 2019a; Yamagishi 1988).

Expanding the strategy space for defenders by allowing a choice between fighting back and running away would enable a more fine-grained analysis of the neural and emotional responses triggered in defenders. Particularly interesting in this regard is **McNaughton & Corr**'s distinction between the anxiolytic-sensitive Behavioral Inhibition System that mediates defensive attention and arousal, and the panicolytic system that mediates fight-flight-freeze responses. It helps to decompose a vigilant defense from the outward anger that defenders may experience when facing the threat of attack (**Andrews, Huddy, Kline, Nam, & Sawyer** [**Andrews et al.**]). Expanding the strategy space with disengagement options would also allow the detection of trait-based differences in threat responding, with some individuals being more likely to protect themselves by fighting and others by withdrawing and disengaging from the relationship. The neuropsychological model sketched in McNaughton & Corr can serve as an excellent starting point for uncovering such individual differences and the model's underlying biology (also see **Paiva, Coelho, Paison, Ribeiro, Almeida, Ferreira-Santos, Marques-Texeira, & Barbosa** [**Paiva et al.**]).

Expanding the strategy space by including disengagement options can have important implications for intergroup attacker-defender conflicts. We agree with **Buckner & Glowacki** and **Fog** that, when individuals within defender groups can flee as an alternative to contributing to collective defense, the typical dynamics we see in intergroup attacker-defender contests may change. Free-rider incentives are typically stronger in attacker compared with defender groups, but such difference disappears when individuals in defender groups can disengage and flee from the group, especially when the anticipated costs of disengagement is lower than the anticipated costs of defense. The mere presence of such disengagement options may also undermine the defender group's cohesion and sense of shared identity, rendering it important for group leaders to create and build group identification and commitments among its members. Fog (also see **Simandan**) discusses this from an evolutionary perspective, suggesting that, when disengagement options are available, defensive warfare also may have given rise to preferences for strong leadership, discipline, punishment institutions, and intolerance of deviants.

Although not mentioned in the commentaries, expanding the strategy space with disengagement options should not be confined to defense. In as much as defenders may have a choice between fighting back and running away as a means to survive attacks, attackers may have a choice between attacking and production to increase wealth (Carter & Anderton 2001; Duffy & Kim 2005; Grossmann & Kim 2002). For example, organizations seeking to increase their profit margins can attempt a hostile takeover, invest in innovative production technologies, or some combination of both. Again, such alternative strategies essentially mean that (groups of) individuals reduce the interdependency

within and/or between groups and forego the benefits of possible cooperation or conflict.

When disengagement options reside within only attackers or defenders, power differences emerge. Attacker threat becomes less pressing, for example, when defenders have solid escape options to complement the resources available for defensive aggression. Accordingly, bargaining and negotiation research showed that having a "Best Alternative to Negotiated Agreement" firms up negotiators, leading them to ask more and concede less (**Halevy**; also see Bazerman & Neale 1985; Carnevale & Pruitt 1992; Giebels et al. 2000; Pinkley 1995). Likewise, studies of public goods provision showed that the threat of punishment is ineffective when participants have outside options available and can thus escape costly sanctions (Gross & De Dreu 2019a; Mulder et al. 2006). In short, when the (quality of) disengagement options are differentially distributed among attackers and defenders, differences in dependency emerge that render the less dependent party more powerful (Barclay & Raihani 2016; Orbell & Dawes 1993; Yamagishi 1988).

Asymmetries in dependency are but one reason for power differences to emerge between attackers and their defenders. Our commentaries raise two other sources of power that are both related to the ability to coerce the antagonist into submission – outnumbering the antagonist and having surplus resources to invest in fighting (**Andrews et al.**; **Buckner & Glowacki**; **Radford, Schindler, & Fawcett** [**Radford et al.**]; **Ridley & Mirville**; **Shnabel & Becker**; **Weisel**). Although differences in coercive capabilities and/or dependencies are theoretically orthogonal to the attacker or defender position, power differences may profoundly influence attack propensity and/or willingness to defend (versus surrender or fleeing) (**Hafer**). In his commentary, **Weisel** provides a generalized form of our basic attacker-defender game, which allows predictions when power differences between attacker and defender emerge and how such power differences should impact behavioral decisions related to attack and defense.

With regard to power differences, **Shnabel & Becker**'s analysis of the psychology of advantaged and disadvantaged groups suggests complex interactions between the attacker versus defender position on the one hand, and the power differential vis-à-vis antagonist on the other. Specifically, disadvantaged groups that may have a latent desire to change the status quo (viz., attacker) are often apathetic, risk-averse, feel inferior, and lack confidence. Advantaged groups who stand to only lose (viz., defenders), in contrast, are more energetic, risk-tolerant, with stronger feelings of deservingness and superiority. History provides ample examples of such society-level dynamics in which the oppressed serve and justify their oppressors, including the Apartheid regime in South Africa, immigrant groups in contemporary Western societies, and enslaved tribal communities at the height of the Roman Empire (also see **Andrews et al.**). We suggest that Shnabel & Becker's important analysis can help explain why power differentials within societies can perpetuate and that disadvantageous groups remain passive and shun challenging the status quo, exactly because of a lack of risk-tolerance, confidence, and feelings of deservingness. From this lens, reinforcing a feeling of inferiority in disadvantageous groups, through, for example, racial or social ideology, can be seen as a means of advantageous groups (viz., defenders) to prevent attackers from developing the psychological prerequisites necessary for challenging the status quo and initiating a conflict. Societal disparities in wealth and power thus can be a source of conflict, but Shnabel & Becker's analysis of advantaged and disadvantaged groups highlight the important point that, next to economic factors, psychological factors need to be met before attacker-defender conflict arises.

### *R2.3. Games of strategy and matching-mismatching of attack and defense*

In section 2 of our target article, we briefly referenced games of strategy that share key properties with the AD-G, including the hide-and-seek game, the matching pennies game, the inspection game, and the best-shot-weakest link game (**Chowdhury**; **Krawczyk**; **Sheretema**). Among these key features that set asymmetric conflicts apart from symmetric conflicts (including the PD-Alt discussed in **Huffmeier & Mazei**, which has multiple pure Nash equilibria) is that attackers optimize their earnings by mismatching their defenders' strategy – compete when the other cooperates, otherwise cooperate – whereas defenders optimize their earnings by matching – compete when the other competes, otherwise cooperate.

Whereas action-reaction tendencies are core to the behavioral study of conflict and conflict resolution (e.g., Axelrod 1984; Carnevale & Pruitt 1992), we have limited insight into matching-mismatching in asymmetric conflicts of attack and defense. **Krawczyk** offers a useful entry to the formal and empirical literature of the general matching pennies game (Goeree et al. 2003; also see Eliaz & Rubinstein 2011; Franke et al. 2013), and **Lopez** provides a compelling discussion of mismatching and matching during coalitional conflicts and tribal raiding in particular. Both commentaries serve as excellent starting points for new research into the question of when and why people (fail to) mismatch during attack, and match during defense. In particular, the observation that mismatching may be more difficult and "counter-intuitive" than matching (Belot et al. 2013; Crawford & Iriberri 2007; Li & Camerer 2019) could explain why defenders not only are faster, but also disproportionately often survive their antagonists' attacks in laboratory experiments (see **Buckner & Glowacki** and sect. R2.4). And it would fit the idea that evolutionary selection has favored ability for matching over mismatching, because failure to match during defense can be more devastating (i.e., foregoing life) than failure to mismatch (i.e., foregoing dinner; Dawkins & Krebs 1979; also see **Hafer**; **Mifune & Simunovic**; **Weisel**).

### *R2.4. Simultaneous versus sequential moves of attack and defense*

The AD-G developed in the target article assumes that antagonists move simultaneously. Several commentaries highlight that, oftentimes, antagonists can or have to move sequentially (**Buckner & Glowacki**; **Lopez**; **Simandan**). In theory, such sequential decision-making in which either attackers or defenders select their strategy before the antagonist does should matter more, strategically and psychologically, when conflict outcomes are probabilistic rather than deterministic and when knowledge about the antagonist's strength is incomplete or imperfect. Under such conditions, attackers may have good reasons to strike first, or in the words of war scholar Von Clausewitz (1832/1984): "Time … is less likely to bring favor to the victor than to the vanquished. … An offensive war requires above all a quick, irresistible decision. … Any kind of interruption, pause, or suspension of activity is inconsistent with the nature of offensive war" (p. 611). It is interesting to note that work reviewed by Buckner & Glowacki (also see Lopez) provides ample counter-examples,

where attackers take their time to carefully design their attack strategy and minimize risk of casualties, and defenders act swiftly (including fleeing the scene). Their observation that such strategic use of time and planning is seen among nonhuman primates as well. Combined with the reproductive fitness functionalities of being a successful attacker (**Becker & Dubbs**; Buckner & Glowacki), this suggests that such strategic timing of attack behavior is adaptive.

Related to the issue of moves and countermoves is whether the attacker-defender contest is operationalized as a one-shot interaction or as a repeated interaction with a shadow of the past and future (**Radford et al.**; **Ridley & Mirville**; **Rusch & Böhm**). In some of our work, discussed in the target article, such ongoing interactions between attackers and defenders have been studied. Results show that attackers "track" their defenders' history of play, form predictions about defenders' likely strength in the next contest round, and adapt accordingly (e.g., De Dreu et al. 2016a; Zhang et al. 2019). This initial work can be extended in two important directions. First, with repeated interactions, there is the possibility of *role shifts*, where defenders who "survived" an attack turn the table and become attackers themselves, forcing their attackers into a defensive position. Radford et al. and Rusch & Böhm highlight how even anticipating such a possibility of role shifts and the concomitant fear of retaliation can already impact the likelihood and forcefulness with which attackers move against their defenders. Such role shifts also explain why defenders sometimes display anger and contempt (see **Andrews et al.**). We expect such approach-related emotions to emerge, especially when role shifts are possible and defenders can counter-attack and retaliate against their (former) aggressors.

The second key extension for the work on repeated attacker-defender contests is to make future fighting power conditional on past success. Indeed, nonhuman predators consume energy and can only repeat the chase a limited number of times until they are too depleted and weak to further attack their prey – predators can afford only a limited number of attacks until starvation becomes a serious possibility. Likewise, prey may successfully ward off initial attacks, but they may lack the resources and strength to ward off subsequent ones. Examples of attackers trying to starve the defenders until the point that they either surrender or are too weak to fight back are also abound in human conflicts. Yet, whereas this dynamic is well-documented and modeled in the literature on nonhuman predator-prey conflicts (**Radford et al.**; **Ridley & Mirville**), the study of human conflicts has largely ignored the dynamic increase or decrease in fighting capacity as a function of past success and failure. New work is needed to understand conflict dynamics when the lure of victory is countered by fear of retaliation and the relief of survival is countered by the threat of renewed attacks. We agree with Radford et al. and Ridley & Mirville that the work on animal conflict can help inform our understanding of human conflict in this regard (and many others).

### *R2.5. Summary and conclusions*

When one party wants a change that is costly to the other side, attack-defense structures emerge in which parties may seek to realize their goals through a range of more or less competitive strategies and tactics. Our basic model of attacker-defender conflicts can be extended in two fundamental ways: (1) by allowing conflict outcomes to be probabilistic rather than deterministic, and (2) by incorporating differences in dependency and coercive capability. To understand strategic choices and tactical maneuvering, it will be useful to incorporate the shadows of the past and future, in which attackers and defenders react to their antagonist's prior moves, or can switch roles and retaliate. Incorporating such structural components would enable an even more fine-grained understanding of asymmetric conflicts within and between groups, including underlying biological, psychological, and sociocultural mechanisms. It also allows us to identify the important factors that predict under which circumstances attack-success increases.

## R3. Framing the game and aligning people to fight

Among the main contributions advanced by the psychological sciences is that humans act on their subjective interpretation of the situation they are in (Halevy et al. 2019; Rauthmann et al. 2014). Whereas we can identify conflict structures as asymmetric with or without a past and a future, and with or without power differences between the antagonists, what matters as much, if not more, is how people "perceive the game" (Balliet et at. 2017; Halevy et al. 2006). Thus, when the structure of the conflict allows for integrative, mutual gains but people perceive it as a winner-takes-all conflict, they fight rather than negotiate and oftentimes "leave money on the table" (De Dreu et al. 2000; Gelfand & Realo 1999; Halevy et al. 2011). Culture, socialization, and perhaps even biological factors condition how people interpret their natural and social surroundings and can, accordingly, profoundly impact their approach to conflict and conflict resolution (Halevy et al. 2011; 2019). In our target article (sects. 3 and 4), we touched upon the possibility that the structure of attacker-defender conflicts may not perfectly map onto the way the conflict, and one's role therein is perceived and enacted. Our commentaries pursue this further and in more detail (**Halevy**; **Pärnamets, Reinero, Pereira, & Van Bavel** [**Pärnamets et al.**]; **Rusch & Böhm**; **Urbanska & Pherson**) with regard to (1) the endogenous emergence of attacker and defender roles, and (2) the sociocultural interventions that frame the goals that groups of people pursue and commit to.

### *R3.1. Endogeneity of attacker and defender roles*

**Hafer** makes a unique contribution to our theoretical outlook by identifying a strategic mechanism that explains role-contingent differences in conflict. She shows how population-wide differences in the ownership of assets emerge as a function of winning symmetric contests (e.g., for unclaimed, new territory), thus creating "haves" and "have-nots." Whereas the haves stand something to lose and wish to defend their wealth (viz., defenders), the have-nots have something to gain, emerging as potential attackers. The intriguing prediction Hafer advances is that the population-wide distribution of defenders dominates that of attackers, something akin to the advantaged and disadvantaged groups addressed in **Shnabel & Becker** and in **Andrews et al.** Crucially, Hafer's analysis can explain the evolved neurobiological responses to attack and defense that we outlined in section 3 of our target article.

Several commentaries draw on evolutionary psychology to propose that males have evolved capacities to fit attack, whereas females are more likely to have evolved capacities to defend (**Becker & Dubbs**; **Lopez**). It would follow that females attack less aggressively than males, yet they defend at least as aggressively, if not more, than males. At present, however, we have no

data to support such possibilities. When we compare the sexes in terms of effort spent in attack-defense contests, we find no significant interactions between sex and role (De Dreu et al. 2019; De Dreu & Giffin 2018). Likewise, in the context of coalitional warfare, it may be that males have an evolved psychology to attack more than females (who have an evolved psychology to contribute to in-group defense [Lopez]). Again, however, we have no data to support such a possibility. In De Dreu et al. (2016a), we were able to compare all-male, all-female and mixed-sex groups but found no differences in neither attack nor defense in a laboratory game setup. However, the study was not designed to examine sex-differences and the sample size was rather small. Intergroup AD-Gs, as proposed in our target article, along with the generalized versions developed in **Weisel**, can help to further elucidate this possibility of socially construed or biologically prepared sex-specific roles in asymmetric conflicts within or between groups of people.

Whereas the formal analysis offered by **Hafer**, and to some extent the evolutionary arguments for possible sex-differences by **Becker & Dubbs** and **Lopez**, purport that clear-cut defender and attacker types emerge, several commentaries emphasize that it is oftentimes unclear who is, or feels, to be an attacker or defender. **Rusch & Böhm** discuss two psychological mechanisms that bias people's perceptions of the conflict and their respective roles therein. Schema-based distrust, in which people unduly fear exploitation by rivaling out-groups, is one such mechanism that Rusch & Böhm suspect may lead people to feel being in a defender position and motivates preemptive aggression of out-groups. In keeping with our target article, we subsume schema-based distrust under the broader header of hostile attribution bias that serves defense and can, as we noted, trigger preemptive strikes even when no actual out-group danger exists. We agree with Rusch & Böhm that being the target of a preemptive strike by a trigger-happy defender may turn otherwise innocent and peace-abiding groups willing to retaliate. In such escalatory spirals of preemptive strikes and retaliatory counter-strikes, both sides may honestly feel being the defender against an unreasonably hostile out-group. Reconstructing who started in which position first or last becomes another psychological tool in the toolbox of conflict parties to motivate future collective action.

The second mechanism discussed in **Rusch & Böhm** is the explicit framing of one's own position as defensive rather than offensive. **Halevy** likewise discusses work on the mental representation of conflict, showing that people often perceive international conflicts as an asymmetric game in which "we" defend and "they" aggress (e.g., Halevy et al. 2006; Plous 1985). Consistent with our argument that being in a defender position mobilizes greater support for the group's cause than being in an attacker position, such explicit framing can help overcome the problem of incentive misalignment present in attacker groups (Halevy; Rusch & Böhm; also see **Simandan; Andrews et al.**). **Pärnamets et al.** suggest that effective leaders may have an intuitive grasp of the malleability of attack-defense dynamics and use rhetoric and propaganda to "frame the game" in terms of defense rather than offense. History provides ample examples of such framing and reframing (see also sect. 4.3 in the target article).

### *R3.2. Group identity and sacralization as incentive alignment strategies*

A key argument developed in our target article is that group defense permits the endogenous emergence of in-group affiliation and identification more than out-group attack. **McLoughlin & Corriveau** take this argument further, using insights from developmental psychology. It is interesting to see that young children's in-group bias is first and foremost oriented towards the positivity of their in-group, driving loyalty and propensity to cooperate with similar others. Only at later age, children develop negative out-group bias as well, showing tendencies to derogate and discriminate against others who are "different."

Such different developmental trajectories underlying early positive in-group bias and later negative out-group bias fit meta-analytic evidence showing that people are more likely to cooperate with in-group members, than to compete against out-group members (Balliet et al. 2014; also see Brewer 1979). We note with **O** that, indeed, the primary functionality of the in-group for young children is safety and protection, fitting the idea that developing a propensity for (in-group) defense early in life and more than for (out-group) attack is adaptive. **Mifune & Simunovic**, likewise, note that defensive motivation more than the desire to aggress and subordinate could be key to the evolved capacity for parochial altruism and in-group bounded cooperation in humans. **Hurlemann & Marsh** offer the possibility that the structurally preserved oxytocinergic system may modulate such parochial altruism aimed at preserving and protecting the in-group, if needed through offensive actions that neutralize the dangers posed by hostile out-groups (viz., preemptive strikes; also see De Dreu et al. 2010; 2011; Ten Velden et al. 2017; Zhang et al. 2019).

While accepting the evidence, some commentaries noted that attacker groups not necessarily lack in-group identification and commitment, or even that in-group identification and commitment among attacker groups can be stronger than in defender groups. **Simandan; O**; and **Fog** all note, for example, that defender groups may be heterogeneous in their perception of out-group threat, or that specific factions within a defender group would suffer more from defeat than others. Such heterogeneity undermines a feeling of shared common fate and concomitant identification with and loyalty to their (defender) group. Vice versa, **Katna & Cheon** note that individuals in attacker groups may, through a process of *identity fusion*, immerse in their group and commit to the point where self-sacrifice is seemingly unavoidable and the only right thing to do.

Although we acknowledge that attacker groups may display strong(er) identification and commitment in some circumstances, we maintain that, *all else constant*, in-group identification and commitment are more likely to endogenously emerge when defending, and exogenous interventions by, for example, group leaders or institutions, are needed more to motivate attack. However, we have only limited evidence for our hypothesis, and herein lies a key target for future research. Such work could explore two possibilities. The first is leader rhetoric (**Pärnamets et al.**), which we discuss in section 4 of the target article. The second is sacralization and moral rigidity, a possibility raised by **Marie**. Sacralization refers to the all-or-nothing valuation of core social obligations, symbols, or natural resources to the extent that these obligations, symbols, and resources become a defining attribute of the in-group's identity and cause. Marie hypothesizes that humans have an evolved capacity to sacralize and reify moral obligations to attract the trust of in-group members, akin to the idea that parochial altruism signals loyalty to the group and leads to potential benefits through direct and indirect reciprocity within one's group (Balliet et al. 2014; Brewer 1979; Yamagishi & Kiyonari 2000).

Some support for **Marie**'s hypothesis derives from Ledgerwood, Liviatan, and Carnevale (2007) who showed, across four studies, that the value placed on material symbols (e.g., a building) depends on commitment to group identity, the extent to which a symbol can be used to represent in-group identity and situational variability in goal strength induced through group-identity affirmation or threat. Thus, property derives value from its capacity to serve as an effective means in the pursuit of group-identity goals. Also consistent with Marie's hypothesis is work showing that individuals negotiating moral as opposed to resource conflicts have stronger win-lose perceptions and therefore are less able to reach mutually beneficial, integrative agreements (Harinck et al. 2000; Harinck & Druckman 2017). It follows that sacralization and the resulting moral rigidity enable groups to, first, overcome possible incentive misalignments within their group through enhanced identification with their in-group. Second, moral rigidity can justify aggressive attacks on neighboring groups in terms of the sacred protection of the in-group's moral legacy and superiority.

### R3.3. Summary and conclusions

Modeling conflict as an asymmetric game of attack and defense provides a lens through which conflict can be analyzed. Compared with symmetric models of conflict, asymmetric conflict models have stronger ecological validity, in that the majority of conflicts between individuals and their groups evolve around the desire to change versus to protect the status quo. Our commentators highlight another reason why conceptualizing conflict as an asymmetric game of attack and defense is important. Asymmetries are not only found in the structure of conflict, but also emerge in the subjective perceptions of one's own role in the conflict. Perceiving oneself as a defender of the in-group and its sacred resources and superior moral stance may be more fitting than perceiving oneself as an attacker of out-groups. Being a defender of the status quo may be more amenable to building and maintaining a positive view of oneself and the in-group than being a proponent of change and revision. This possibility could explain why leader rhetoric and propaganda emphasize the moral superiority and deservingness of the in-group along with the moral inferiority and threat inherent in rivaling out-groups. As we argued the functionality of such self-serving distortions is, first and foremost, reducing the incentive-alignment problem (making costly contributions) along with the coordination failure (organizing collective action at the right time and with the proper force) that groups suffer from attacking out-groups more than when defending the in-group against out-group threat. Exploring the psychological mechanism that allows individuals and groups to frame themselves as defenders and legitimize their actions may help us understand when and why conflicts arise and persist.

## R4. Transforming the game: Solving attacker-defender conflict

Although our main goal was to highlight and develop asymmetric conflict theory, an important application of conflict theory is conflict resolution and dispute settlement. Our target article showed that attacker-defender conflict may require different interventions than symmetric conflicts, precisely because of the distinctly different roles and goals that attackers and defenders have for starting the conflict and continuing it. We focused on third-party interventions aimed at attackers, arguing that if third-party interventions can either improve the status quo or tax the possible spoils of war, attackers should be less motivated to compete and more motivated to accept the status quo. **Urbanska & Pherson** discuss the role of authority legitimacy, rightfully noting that outside interventions sometimes backfire when performed by third parties who lack the legitimate authority to do so. **Halevy** invokes negotiation theory, and **Cernadas Curotto**, **Halperin**, **Sander**, **& Klimecki (Cernadas Curotto et al.)** consider emotion regulation as additional means for conflict resolution. These we discuss in some detail.

### R4.1. Negotiating settlements

Negotiation, with or without assistance from uninvolved third parties, is a tried-and-true technique for resolving conflict and reaching lasting agreements (Kelman 2006; Lax & Sebenius 1986; Pruitt & Rubin 1986). Using our attacker-defender game as a backdrop, **Halevy** develops important insights for motivating attackers and defenders to give up fighting and to "come to the table" to negotiate an agreement. For such negotiations to work, Halevy rightfully notes that the game needs to be transformed into a coordination game in which both sides can actually win something. In a similar vein, **Shnabel & Becker** rightfully point out that a change in the status quo, desired by attackers, does not necessarily have to result in a loss for defenders. To defenders, a win could take the form of an increased sense of security; for attackers, it could take the form of an improved status quo. Negotiation scholars have developed various techniques for creating such "integrative potential," including (1) increasing the number of issues that is part of the negotiation; (2) decomposing a few broad issues into multiple smaller ones; (3) considering issues in terms of underlying needs (e.g., security, prosperity); and (4) considering issues and their implications for need fulfillment, in combination rather than in isolation (Lax & Sebenius 1986; Pruitt 1981; Raiffa 1982; Walton & McKersie 1965). We agree with Halevy that negotiation theory and research offer extant possibilities for constructive resolution of attacker-defender conflicts within and between groups of people. Further, the insight that social games are often differently perceived and construed on the psychological level (as touched upon by Halevy; **Pärnamets et al.**; **Rusch & Böhm**; Shnabel & Becker; **Urbanska & Pherson**) points to important intervention possibilities.

**Halevy** also suggests that negotiation theory offers insights into how attackers and defenders can be motivated to initiate negotiations. An important additional insight here derives from so-called *readiness theory* (Pruitt 2007; Zartman 1989; 2000). In brief, the idea is that antagonists shift from fighting to negotiation when there is (1) a mutually hurting stalemate in which continuation of the conflict is exceedingly costly (i.e., being stuck in a "bad" equilibrium), and (2) an optimistic belief that the other side is willing to lower its aspirations and able to make concessions. For example, the 1998 peace agreement between the Irish Republican Army (IRA) and the United Kingdom (UK) ended a bloody and mutually hurting conflict – the Troubles – over the independency of Northern Ireland. Pruitt (2007) attributes the outcome to (a) IRA and British discouragement about the likelihood of a military victory, (b) pressure from both sides' allies and constituencies, and (c) growing optimism about the success of negotiation. In terms of our analysis, the Troubles can be conceived of a basic attacker-defender game between the revisionist IRA and the non-revisionist UK government. The lasting peace that was negotiated more than

two decades ago indeed suggests, that negotiation can be instrumental in resolving attacker-defender conflicts. Readiness theory provides a good starting point to analyze and predict when and why attackers and defenders initiate negotiations as a means to resolve their differences.

### R4.2. Regulating emotions

Inherent in readiness theory and critical to get negotiations started is an element of hope that future waste can be prevented, and optimism about creating an end to the conflict (Bar-Tal 2001; Pliskin & Halperin 2016; Pruitt 2007). Hopelessness and concomitant apathy may be, indeed, among the key emotional states that characterize disadvantaged groups in society (**Shnabel & Becker**). Optimism requires the belief that the other can change (viz., malleability; Halperin et al. 2011). Thus, to get negotiations started and to seek constructive rather than violent resolution of conflict, interventions may target the antagonist's hope and optimism.

Work summarized by **Cernadas Curotto et al.** shows that this can be done and indeed contributes to constructive conflict resolution. For example, Cernadas Curotto et al. draw on the idea that people are motivated to feel certain ways, and we agree that defenders may (i) have different emotional preferences than attackers, because certain emotions (ii) are instrumental to the antagonist's goals in the conflict. **Sheretema** discusses how such emotional states and preferences like guilt and inequity aversion, on the one hand, and anger and regret aversion, on the other hand, can lead to substantial deviations from what rational selfish agents in attacker-defender conflicts should do. Indeed, in recounting his experiences as a mediator in the Balkan conflicts, Holbrooke (1999) describes a good example of such instrumental use of emotions: "Karadzic…said that our draft proposal was unacceptable. Suddenly, Mladic erupted. Pushing to the center of the circle, he began a long, emotional diatribe. … This was the intimidating style he had used with the Dutch commander at Srebrenica, with Janvier, and with so many others. He gave off a scent of danger. … I did not know if his rage was real or feigned, but this was the genuine Mladic, the one who could unleash a murderous rampage" (pp. 150–51). Cernadas Curotto et al. discuss several interventions to change emotions and emotion-based preferences, including reappraisal training and compassion training. Compassion training, in particular, may enable attackers to inhibit their willingness to change the status quo through violence and contribute to a de-escalatory move that allows both the attacker and the defender to negotiate rather than fight.

### R4.3. Summary and conclusions

Asymmetric conflicts between attackers and defenders may not only be more frequent than the widely studied symmetric conflicts, but they may also offer and require different measures and interventions for conflict resolution and peace settlement. Next to the economic interventions we discussed in our target article, research and theory on negotiation, readiness, and emotion regulation offer interventions for conflict resolutions and suggest important pathways to peace.

## R5. Conclusion

The conflicts that humans create and fight within and between groups can be meaningfully modeled as games of strategy. Grounded in the observation that emerging conflicts are more often between those who seek change and revision of the status quo, and those who seek to maintain and protect the status quo, we proposed to consider attacker-defender conflicts in more detail.

Our framework, along with the commentaries on our target article, largely focused on human conflict and the neuropsychological and sociocultural mechanisms that operate during attack and defense. The commentaries refined and added insights about the structural features of asymmetric conflict, the strategies people choose, and the tactical maneuvering that can take place, along with key moderators of group identification and possibilities for conflict resolution.

Whereas the study of human conflict largely neglected asymmetric conflicts between attackers and defenders, scholars in biology have long recognized the distinct dynamics between (group-hunting) predators and (herds of) prey. Without denying the possibility of unique psychological and cultural capabilities of the human species, we agree with **Radford et al.** and **Ridley & Mirville** that integrating the study of animal conflict with that of human conflict can be mutually beneficial and fruitful. Among other things, such integration can shed light on the long-term selection pressures emanating from asymmetric conflicts between attackers and defenders (**Hafer; Mifune & Simunovic**), including the possible group-selection pressures on the emergence of the (human) propensity for cooperation, indirect reciprocity, and parochial altruism (viz., Bowles & Gintis 2011). Ultimately, such integration should enable a biologically tractable, ecologically valid, and psychologically plausible theory of conflict and cooperation within and between groups that is amenable to interventions for constructive conflict resolution and reduced suffering.

**Acknowledgments.** Preparation of this response article was facilitated by a visiting professorship grant from Sapienza University Rome to CKWDD and has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (AdG Agreement No. 785635) to CKWDD, and a Netherlands Science Foundation VENI Grant (016.Veni.195.078) to JG. The authors have no conflicts to declare.

## References

**[The letters "a" and "r" before author's initials stand for target article and response references, respectively]**

Aaldering H., Greer L. L., Van Kleef G. A., & De Dreu C. K. W. (2013) Interest (mis) alignments in representative negotiations: Do pro-social agents fuel or reduce intergroup conflict? *Organizational Behavior and Human Decision Processes* **120**(2):240–50. [NH]

Aaldering H., Ten Velden F. S., Van Kleef G. A., & De Dreu C. K. W. (2018) Parochial cooperation in intergroup conflict is reduced when it harms out-groups. *Journal of Personality and Social Psychology* **114**(6):909–923. [aCKWDD]

Abbink K. (2012) Laboratory experiments on conflict. In: *The Oxford handbook of the economics of peace and conflict*, ed. M. R. Garfinkel & S. Skaperdas, pp. 532–53. Oxford University Press. [aCKWDD]

Abbink K., Brandts J., Herrmann B. & Orzen H. (2010) Inter-group conflict and intra-group punishment in an experimental contest game. *American Economic Review* **100**:420–47. [aCKWDD]

Abbink K., Brandts J., Hermann B. & Orzen H. (2012) Parochial altruism in inter-group conflicts. *Economic Letters* **117**:45–48. [aCKWDD]

Abbink K. & de Haan T. (2014) Trust on the brink of Armageddon: The first-strike game. *European Economic Review* **67**:190–96. Available at: https://doi.org/10.1016/j.euroecorev.2014.01.009. [aCKWDD, HR]

Abele S., Stasser G. & Chartier T. (2010) Conflict and coordination in the provision of public goods: A conceptual analysis of continuous and step-level games. *Personality and Social Psychology Review* **14**:385–401. [aCKWDD]

Adams J. S. (1966) Inequity in social exchange. *Advances in Experimental Social Psychology* **2**:267–99. Available at: https://doi.org/10.1016/S0065-2601(08)60108-2. [KU]

Adelman H. M. & Maatsch J. L. (1956) Learning and extinction based upon frustration, food reward, and exploratory tendency. *Journal of Experimental Psychology* **52**:311–15. [NMcN]

Aktipis A., Cronk L., Alcock J., Ayers J. D., Baciu C., Balliet D., Boddy A. M., Curry O. S., Krems J. A., Muñoz A., Sullivan D., Sznycer D., Wilkinson G. S. & Winfrey P. (2018) Understanding cooperation through fitness interdependence. *Nature Human Behaviour* **2**(7):429–31. https://doi.org/10.1038/s41562-018-0378-4. [AM]

Albert D. J., Walsh M. L. & Jonik R. H. (1993) Aggression in humans: What is its biological foundation? *Neuroscience and Biobehavioral Reviews* **17**:405–25. [aCKWDD]

Aldrich J. H., Gelpi C., Feaver P., Reifler J. & Sharp K. T. (2006) Foreign policy and the electoral connection. *Annual Review of Political Science* **9**:477–502. [TMA]

Alexander R. (1987) *The biology of moral systems (foundations of human behavior)*. Aldine de Gruyter. [AM]

Allen M. W., Bettinger R. L., Codding B. F., Jones T. L. & Schwitalla A. W. (2016) Resource scarcity drives lethal attack among prehistoric hunter-gatherers in central California. *Proceedings of the National Academy of Sciences USA* **43**:12120–25. [aCKWDD]

Almeida P. R., Seixas M. J., Ferreira-Santos F., Vieira J. B., Paiva T. O., Moreira P. S. & Costa P. (2015) Empathic, moral and antisocial outcomes associated with distinct components of psychopathy in healthy individuals: A triarchic model approach. *Personality and Individual Differences*, **85**:205–11. Available at: https://doi.org/10.1016/j.paid.2015.05.012. [TOP]

Altemeyer B. (1988) *Enemies of freedom: Understanding right-wing authoritarianism*. Jossey-Bass. [PP]

Althaus S. L. & Largio D. M. (2004) When Osama became Saddam: Origins and consequences of the change in America's public enemy# 1. *PS: Political Science & Politics* **37** (4):795–99. [TMA]

Altman N. (2008) On suicide bombing. *International Journal of Applied Psychoanalytic Studies* **5**(1):51–67. [DK]

Alvard M. S. & Nolin D. A. (2002) Rousseau's whale hunt? Coordination among big-game hunters. *Current Anthropology* **43**:533–59. [aCKWDD]

Andreoni J. (1995) Warm-glow versus cold-prickle – The effects of positive and negative framing on cooperation in experiments. *Quarterly Journal of Economics* **110**:1–21. [aCKWDD]

Arce D. G., Kovenock D. & Roberson B. (2012) Weakest-link attacker-defender games with multiple attack technologies. *Naval Research Logistics (NRL)* **59**(6):457–69. [SMC]

Archer J. (2004) Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology* **8**:291–322. [DVB]

Arnsten A. F. T. (2009) Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience* **10**(6):410–22. Available at: https://doi.org/10.1038/nrn2648. [PCC]

Aron A. R., Fletcher P. C., Bullmore E. T., Sahakian B. J. & Robbins T. W. (2003) Stop-signal inhibition disrupted by damage to the right inferior frontal gyrus in humans. *Nature Neuroscience* **6**:115–16. [aCKWDD]

Aron A. R., Robbins T. W. & Poldrack R. A. (2014) Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences* **18**:177–85. [aCKWDD]

Arreguin-Toft I. (2005) How the weak win wars. Retrieved August 8, 2016, from http://www.cambridge.org/gi/academic/subjects/politics-international-relations/international-relations-and-international-organisations/how-weak-win-wars-theory-asymmetric-conflict. [ACL]

Arseneau-Robar T. J. M., Müller E., Taucher A. L., van Schaik C., Bshary R. & Willems E. P. (2018) Male monkeys use punishment and coercion to de-escalate costly intergroup fights. *Proceedings of the Royal Society B: Biological Sciences* **285**:20172323. [ANR]

Arseneau-Robar T. J. M., Taucher A. L., Müller E., van Schaik C., Bshary R. & Willems E. P. (2016) Female monkeys use both the carrot and the stick to promote male participation in intergroup fights. *Proceedings of the Royal Society B: Biological Sciences* **283**:20161817. [ANR]

Ashby F. G., Isen A. M. & Turken A. U. (1999) A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review* **106**:529–50. [aCKWDD]

Atran S. (2010) *Talking to the enemy: Violent extremism, sacred values, and what it means to be human*. Penguin. [AM]

Atran S. (2016) The devoted actor. Unconditional commitment and intractable conflict across cultures. *Current Anthropology* **57**(Suppl. 13):S192–S203. [AM]

Atran S. & Ginges J. (2012) Religious and sacred imperatives in human conflict. *Science* **336**:855–57. [aCKWDD]

Avenhaus R., Canty M., Kilgour D. M., von Stengel B. & Zamir S. (1996) Inspection games in arms control. *European Journal of Operational Research* **90**:383–94. [aCKWDD]

Axelrod R. (1984) *The evolution of cooperation*. Penguin. [arCKWDD]

Babcock L. & Loewenstein G. F. (1997) Explaining bargaining impasse: The role of self-serving bias. *Journal of Economic Perspectives* **11**:109–26. [aCKWDD]

Bacharach S. B. & Lawler E. J. (1981) *Bargaining: Power, politics, and outcomes*. Jossey-Bass. [arCKWDD]

Back E. (2013) Position toward the status quo: Explaining differences in intergroup perception between left- and right-wing affiliates. *Journal of Applied Social Psychology* **43**:2073–82. [aCKWDD]

Baik K. H., Kim I. G. & Na S. (2001) Bidding for a group-specific public-good prize. *Journal of Public Economics*, **82**(3):415–29. [SMC]

Baker Jr. M. D. & Maner J. K. (2008) Risk-taking as a situationally sensitive male mating strategy. *Evolution and Human Behavior* **29**:391–95. [DVB]

Balliet D., Tybur J. M. & Van Lange P. A. M. (2017) Functional interdependency theory: An evolutionary account of social situations. *Personality and Social Psychology Review* **21**:361–88. [rCKWDD]

Balliet D. & Van Lange P. A. M. (2013) Trust, punishment and cooperation across 18 societies: A meta-analysis. *Perspectives on Psychological Science* **8**:363–79. [aCKWDD]

Balliet D. P., Wu J. & De Dreu C. K. W. (2014) In-group favoritism and cooperation: A meta-analysis. *Psychological Bulletin* **140**(6):1556–81. [arCKWDD, NM]

Baray G., Postmes T. & Jetten J. (2009) When I equals we: Exploring the relation between social and personal identity of extreme right-wing political party members. *British Journal of Social Psychology*, **48**(4):625–47. [DK]

Barclay P. & Raihani N. (2016) Partner choice versus punishment in human Prisoner's Dilemmas. *Evolution and Human Behavior* **37**:263–71. [rCKWDD]

Barclay P. & Van Vugt M. (2015) The evolutionary psychology of human prosociality: Adaptations, byproducts, and mistakes. In Schroeder D. A. & Graziano W. G (Eds), *The Oxford handbook of prosocial behavior*, pp. 37–60. Oxford University Press. [DK]

Barclay P. & Willer R. (2007) Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences* **274**(1610):749–53. [AM]

Bar-Hillel M. (2015) Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science* **10**:19–433. [aCKWDD]

Baron J. (1994) Nonconsequentialist decisions. *Behavioral and Brain Sciences* **17**:1–10. [aCKWDD]

Baron J. & Spranca M. (1997) Protected values. *Organizational Behavior and Human Decision Processes* **70**:1–16. [AM]

Bar-Tal D. (2001) Why does fear override hope in societies engulfed by intractable conflict, as it does in the Israeli society? *Political Psychology* **22**:601–27. [rCKWDD]

Bar-Tal D. (2013) *Intractable conflicts: Socio-psychological foundations and dynamics*. Cambridge University Press. [PCC]

Bar-Tal D., Halperin E. & De Rivera J. (2007) Collective emotions in conflict situations: Societal implications. *Journal of Social Issues* **63**(2):441–60. Available at: https://doi.org/10.1111/j.1540-4560.2007.00518.x. [PCC]

Bartra O., McGuire J. T. & Kable J. W. (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* **76**:412–27. [aCKWDD]

Baskin-Sommers A. R., Wallace J. F., MacCoon D. G., Curtin J. J. & Newman J. P. (2010) Clarifying the factors that undermine behavioral inhibition system functioning in psychopathy. *Personality Disorders: Theory, Research, and Treatment*, **1**(4):203–17. Available at: https://doi.org/10.1037/a0018950. [TOP]

Batchelor T. P. & Briffa M. (2011) Fight tactics in wood ants: individuals in smaller groups fight harder but die faster. *Proceedings of the Royal Society B: Biological Sciences* **278**:3243–50. [ANR]

Batson C. (1998) Altruism. In: *Handbook of social psychology*, ed. G. Lindzey, E. Aronson. Wiley. [aCKWDD]

Battigalli P. & Dufwenberg M. (2009) Dynamic psychological games. *Journal of Economic Inquiry* **144**:1–35. [aCKWDD]

Baumard N., André J.-B. & Sperber D. (2013) A mutualistic theory of morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences* **36**:59–122. [AM]

Baumgartner T., Heinrichs M., Vonlanthen A., Fischbacher U. & Fehr E. (2008) Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* **58**(4):639–50. Available at: https://www.cell.com/neuron/fulltext/S0896-6273(08)00327-9?code=-cell-site. [RH]

Baye M. R., Kovenock D. & De Vries C. G. (1996) The all-pay auction with complete information. *Economic Theory* **8**(2):291–305. [SMC]

Baye M. R., Kovenock D. & De Vries C. G. (2012) Contests with rank-order spillovers. *Economic Theory* **51**(2):315–50. [SMC, RMS]

Bazerman M. H., Curhan J. R., Moore D. A. & Valley K. L. (2000) Negotiation. *Annual Review of Psychology* **51**:279–314. [aCKWDD]

Bazerman M. H., Magliozzi T. & Neale M. A. (1985) Integrative bargaining in a competitive market. *Organizational Behavior and Human Decision Processes* **35**:294–313. [rCKWDD]

Becker D. V. (2017) Facial gender interferes with decisions about facial expressions of anger and happiness. *Journal of Experimental Psychology: General* **146**(4):457–61. [DVB]

Becker D. V., Anderson U.S., Neuberg S. L., Maner J. K., Shapiro J. R., Ackerman J. M., Schaller M. & Kenrick D. T. (2010) More memory bang for the attentional buck: Self-protection goals enhance encoding efficiency for potentially threatening males. *Social Psychological and Personality Science* **1**:182–89. [DVB]

Becker D. V., Mortensen C. R., Ackerman J. M., Shapiro J. R., Anderson U. S., Sasaki T., Maner J. K., Neuberg S. L. & Kenrick D. T. (2011) Signal detection on the battlefield:

Priming self-protection vs. revenge-mindedness differentially modulates the detection of enemies and allies. *PLoS One* **6**(9):e23929. doi:10.1371/journal.pone.0023929. [DVB, DK]

Becker D. V., Mortensen C. R., Anderson U. S. & Sasaki T. (2014) Out of sight but not out of mind: Memory scanning is attuned to threatening faces. *Evolutionary Psychology* **12**(5):878–89. [DVB]

Becker D. V. & Srinivasan N. S. (2014) The vividness of the happy face. *Current Directions in Psychological Science* **23**:189–94. [DVB]

Becker J. C., Kraus M. W. & Rheinschmidt-Same M. (2017) Cultural expressions of social class and their implications for group-related beliefs and behaviors. *Journal of Social Issues* **73**:158–74. [NS]

Beebe J., Qiaoan R., Endara M. A. & Wysocki T. (2015) Moral objectivism in cross-cultural perspective. *Journal of Cognition and Culture* **15**(3/4):386–401. Available at: https://doi.org/10.1163/15685373-12342157. [AM]

Bélanger J. J., Caouette J., Sharvit K. & Dugas M. (2014) The psychology of martyrdom: Making the ultimate sacrifice in the name of a cause. *Journal of Personality and Social Psychology* **107**(3):494–515. [DK]

Belot M., Crawford V. P. & Heyes C. (2013) Players of Matching Pennies automatically imitate opponents' gestures against strong incentives. *Proceedings of the National Academy of Sciences USA* **110**(8):2763–68. [rCKWDD, MWK]

Bennett M., Barrett M., Karakozov R., Kipiani G., Lyons E., Pavlenko V. & Riazanova T. (2004) Young children's evaluations of the ingroup and of outgroups: A multinational study. *Social Development* **13**(1):124–41. doi: 10.1046/j.1467-9507.2004.00260.x. [NMcL]

Berkowitz L. & Harmon-Jones E. (2004) Toward an understanding of the determinants of anger. *Emotion* **4**(2):107. [TMA]

Bernard S. (2012) Cohesion from conflict: Does intergroup conflict motivate intragroup norm enforcement and support for centralized leadership? *Social Psychology Quarterly* **75**:107–30. [aCKWDD]

Bernhard H., Fischbacher U. & Fehr E. (2006) Parochial altruism in humans. *Nature* **442**(7105):912–15. Available at: https://www.nature.com/articles/nature04981. [aCKWDD, RH, NM]

Biddle S. (2006) *Military power: Explaining victory and defeat in modern battle*. Princeton University Press. [ACL]

Birnbaum D., Deeb I., Segall G., Ben-Eliyahu A. & Diesendruck G. (2010) The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development* **81**(3):757–77. doi: 10.1111/j.1467-8624.2010.01432.x. [NMcL]

Bishop D. T., Cannings C. & Maynard Smith J. (1978) The war of attrition with random rewards. *Journal of Theoretical Biology* **3**:377–88. [CH]

Bizumic B. (2019) *Ethnocentrism: Integrated perspectives*. London: Routledge. [DS]

Blair R. J. (2006) Subcortical brain systems in psychopathy. In: *Handbook of psychopathy*, ed. C. J. Patrick, 1st edition, pp. 296–312. Guilford Press. [TOP]

Blair R. J. R. (2007) The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences* **11**(9):387–92. Available at: https://doi.org/10.1016/j.tics.2007.07.003. [TOP]

Blair R. J. R., Meffert H., Hwang S. & White S. F. (2018) Psychopathy and brain function: Insights from neuroimaging research. In: *Handbook of psychopathy*, ed. C. J. Patrick, 2nd ed., pp. 401–21. Guilford Press. [TOP]

Blattman C. & Miguel E. (2010) Civil war. *Journal of Economic Literature* **48**:3–57. [aCKWDD]

Bobo L. & Hutchins V. L. (1996) Perceptions of racial group competition: Extending Blumer's theory of group position to a multiracial social context. *American Sociological Review* **61**:951–72. [aCKWDD]

Boehm C. (2009) *Hierarchy in the forest*. Harvard University Press. [aCKWDD]

Boehm C. (2012) Ancestral hierarchy and conflict. *Science* **336**(6083):844–47. [aCKWDD]

Böhm R. (2016) Intuitive participation in aggressive intergroup conflict: Evidence of weak versus strong parochial altruism. *Frontiers in Psychology* **7**:1535–38. [DK]

Böhm R., Rusch H. & Güreck O. (2016) What makes people go to war? Defensive intentions motivate retaliatory and preemptive intergroup aggression. *Evolution and Human Behavior* **37**(1):29–34. Available at: https://doi.org/10.1016/j.evolhumbehav.2015.06.005. [TMA, WB, aCKWDD, HR]

Böhm R., Thielmann I. & Hilbig B. E. (2018) The brighter the light, the deeper the shadow: Morality also fuels aggression, conflict, and violence. *Behavioral and Brain Sciences* **41**:e98. [aCKWDD]

Bohnet I., Greig F., Herrmann B. & Zeckhauser R. (2008) Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review* **98**:294–310. [aCKWDD]

Bolton G. E. & Ockenfels A. (2000) ERC: A theory of equity, reciprocity, and competition. *American Economic Review* **90**:166–93. [aCKWDD]

Bonanni R., Valsecchi P. & Natoli E. (2010) Pattern of individual participation and cheating in conflicts between groups of free-ranging dogs. *Animal Behaviour* **79**:957–68. [ARR]

Bondü R. & Richter P. (2016) Interrelations of justice, rejection, provocation, and moral disgust sensitivity and their links with the hostile attribution bias, trait anger, and aggression. *Frontiers in Psychology* **7**:795–810. [DK]

Book A. S. & Quinsey V. L. (2004) Psychopaths: Cheaters or warrior-hawks? *Personality and Individual Differences* **36**:33–45. [aCKWDD]

Boot N. C., Baas M., Van Gaal S., Cools R. & De Dreu C. K. W. (2017) Creative cognition and dopaminergic modulation of fronto-striatal networks: Integrative review and research agenda. *Neuroscience and Biobehavioral Reviews* **78**:13–23. [aCKWDD]

Bornstein G. (1992) The free-rider problem in intergroup conflicts over step-level and continuous public goods. *Journal of Personality and Social Psychology* **62**(4):597–606. [NH]

Bornstein G. (2003) Intergroup conflict: Individual, group, and collective interests. *Personality and Social Psychology Review* **7**(2):129–45. [aCKWDD, OW]

Bornstein G., Budescu D. & Zamir S. (1997) Cooperation in intergroup, N-person, and two-person games of chicken. *Journal of Conflict Resolution* **41**:384–406. [aCKWDD]

Bornstein G. & Gilula Z. (2003) Between-group communication and conflict resolution in assurance and chicken games. *Journal of Conflict Resolution* **47**:326–39. [aCKWDD]

Bornstein G., Gneezy U. & Nagel R. (2002) The effect of intergroup competition on group coordination: An experimental study. *Games and Economic Behavior* **41**:1–25. [aCKWDD]

Bornstein G., Kugler T. & Zamir S. (2005) One team must win, the other need only not lose: An experimental study of an asymmetric participation game. *Journal of Behavioral Decision Making* **18**:111–23. [aCKWDD]

Bornstein G., Mindelgrin D. & Rutte C. (1996) The effects of within-group communication on group decision and individual choice in the assurance and chicken games. *Journal of Conflict Resolution* **40**:486–501. [aCKWDD]

Bornstein G. & Weisel O. (2010) Punishment, cooperation, and cheater detection in "noisy" social exchange. *Games* **1**:18–33. [aCKWDD]

Bottom W. P. & Studt A. (1993) Framing effects and the distributive aspect of integrative bargaining. *Organizational Behavior and Human Decision Processes* **56**:459–74. [aCKWDD]

Bouwmeester S., Verkoeijen P. P. J. L., Aczel B., Barbosa F., Begue L., Branas-Garza P., Chmura T. G. H., Cornelissen G., Dossing F. S., Espin A. M., Evans A. M., Ferreira-Santos F., Fiedler S., Flegr J., Ghaffari M., Glockner A., Goeschl T., Guo L., Hauser O. P., Hernan-Gonzalez R., Herrero A., Horne Z., Houdek P., Johannesson M., Koppel L., Kujal P., Laine T., Lohse J., Martins E. C., Mauro C., Mischkowski D., Mukherjee S., Myrseth R., Navarro-Martinez D., Neal T. M. S., Novakova J., Paga R., Paiva T. O., Palfi B., Piovesan M., Rahal R. M., Salomon E., Srinivasan N., Srivastava A., Szaszi B., Szollosi A., Thor K. O., Tinghog G., Trueblood J. S., Van Bavel J. J., van 't Veer A. E., Vastfjall D., Warner M., Wengstrom E., Wills J. & Wollbrant C. E. (2017) Registered replication report: Rand, Greene, and Nowak, 2012. *Perspectives on Psychological Science* **12**:527–42. [aCKWDD]

Bowles H. R., Babcock L. & Lai L. (2007) Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes* **103**(1):84–103. [NH]

Bowles S. (2009) Did warfare amongst ancestral hunter and gatherers affect the evolution of social behaviors? *Science* **324**:1293–98. [aCKWDD]

Bowles S. & Gintis H. (2011) *A cooperative species: Human reciprocity and its evolution*. Princeton University Press. [arCKWDD]

Boyer P., Firat B. & Van Leeuwen F. (2015) Safety, threat, and stress in intergroup relations: A Coalitional Index Model. *Perspectives on Psychological Science* **10**:434–50. [aCKWDD]

Boyer P. & Liénard P. (2006) Why ritualized behavior? Precaution systems and action parsing in developmental, pathological and cultural rituals. *Behavioral and Brain Sciences* **29**:1–56 [aCKWDD]

Brady W. J., Van Bavel J. J., Jost J. & Wills J. A. (2019) An ideological asymmetry in the diffusion of moralized content among political elites. *Journal of Experimental Psychology General*. Available at: https://psycnet.apa.org/doiLanding?doi=10.1037%2Fxge0000532. [PP]

Braver T. S. (2012) The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences* **16**:106–13. [aCKWDD]

Brewer M. B. (1979) In-group bias in the minimal intergroup situation – Cognitive-motivational analysis. *Psychological Bulletin* **86**:307–24. [rCKWDD]

Brewer M. B. (1999) The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues* **55**(3):429–44. doi: 10.1111/0022-4537.00126. [NMcL]

Brewer M. B. & Caporael L. R. (2006) An evolutionary perspective on social identity: Revisiting groups. In: *Evolution and social psychology*, ed. M. Schaller, J. Simpson & D. Kenrick, pp. 143–61. Psychology Press. [NMcL]

Brewer M. B. & Kramer R. M. (1986) Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology* **50**:543–49. [aCKWDD]

Brodie III E. D., & Brodie Jr. E. D. (1999) Predator-prey arms races. *BioScience* **47**:557–68. [aCKWDD]

Brooks D. J. & Valentino B. A. (2011) A war of one's own: Understanding the gender gap in support for war. *Public Opinion Quarterly* **75**(2):270–86. Available at: https://doi.org/10.1093/poq/nfr005. [ACL]

Brown J. S. (1948) Gradients of approach and avoidance responses and their relation to level of motivation. *Journal of Comparative and Physiological Psychology* **41**:450–65. [NMcN]

Bruintjes R., Lynton-Jenkins J., Jones J. W. & Radford A. N. (2016) Out-group threat promotes within-group affiliation in a cooperative fish. *The American Naturalist* **187**:274–82. [ANR]

Brundage J. A. (1962) *The Crusades: A documentary survey*. Marquette University Press. [aCKWDD]

Brunnschweiler C. N. & Bulte E. H. (2009) Natural resources and violent conflict: Resource abundance, dependency, and the onset of civil war. *Oxford Economic Papers* **61**:651–74. [aCKWDD]

Buhaug H. & Rod J. (2006) Local determinants of African civil wars 1970–2001. *Political Geography* **25**:315–335. [aCKWDD]

Burke M., Hsiang S. M. & Miguel E. (2015) Global non-linear effect of temperature on economic production. *Nature* **527**:235–239. [aCKWDD]

Buss D. M. (1989) Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences* **12**:1–49. [DVB]

Buttelmann D. & Böhm R. (2014) The ontogeny of the motivation that underlies in-group bias. *Psychological Science* **25**:921–27. doi: 10.1177/0956797613516802. [NMcL]

Calo-Blanco A., Kovářík J., Mengel F. & Romero J. G. (2017) Natural disasters and indicators of social cohesion. *PLoS One* **12**:e0176885–13. [aCKWDD]

Camerer C. F. (2003) *Behavioral game theory*. Princeton University Press. [aCKWDD]

Campbell A. (1999) Staying alive: Evolution, culture, and women's intrasexual aggression. *Behavioral and Brain Sciences* **22**:203–52. [DVB]

Campbell D. T. (1972) On the genetics of altruism and the counter-hedonic components in human culture. *Journal of Social Issues* **28**:21–37. [aCKWDD]

Campbell D. T. (1975) On the conflicts between biological and social evolution and between psychology and moral tradition. *American Psychologist* **30**:1103. [aCKWDD]

Cant M. A., Otali E. & Mwanguhya F. (2002) Fighting and mating between groups in a cooperatively breeding mammal, the banded mongoose. *Ethology* **108**:541–55. [ARR]

Carneiro R. (1981) The chiefdom as precursor of the state. In: *The transition to statehood in the New World*, ed. G. Jones & R. Krautz. pp. 37–79. Cambridge University Press. [aCKWDD]

Carnevale P. J. (1986) Strategic choice in mediation. *Negotiation Journal* **2**:41–56. [aCKWDD]

Carnevale P. J. & Pruitt D. G. (1992) Negotiation and mediation. *Annual Review of Psychology* **43**:531–82. [arCKWDD]

Carpenter M., Nagell K. & Tomasello M. (1998) Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development* **63**(4):i–vi, 1–143. [NMcL]

Carter J. R. & Anderton C. H. (2001) An experimental test of a predator–prey model of appropriation. *Journal of Economic Behavior & Organization* **45**(1): 83–97. [arCKWDD, MWK]

Carver C. S. 2004. Negative affects deriving from the behavioral approach system. *Emotion* **4**(1):3–22. [TMA]

Carver C. S., Johnson S. L. & Joormann J. (2008) Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: What depression has in common with impulsive aggression. *Psychological Bulletin* **134**(6):912–43. doi:10.1037/a0013740. [NMcN]

Carver C. S. & White T. L. (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology* **67**:319–33. [aCKWDD]

Cernadas Curotto P., Sander D., Halperin E. & Klimecki O. (in preparation) The impact of compassion and emotion regulation training on conflict resolution. [PCC]

Chagnon N. A. (1988) Life histories, blood revenge, and warfare in a tribal population. *Science* **239**(4843):985–92. [WB]

Chambers J. R., Baron R. S. & Inman M. L. (2006) Misperceptions in intergroup conflict. *Psychological Science* **17**:38–45. [aCKWDD]

Chas A., Betancor V., Delgado N. & Rodríguez-Pérez A. (2018) Children consider their own groups to be more human than other social groups: Evidence from indirect and direct measures. *Social Psychology* **49**(3):125–34. [NMcL]

Chen E. E., Corriveau K. H., Lai V. K. W., Poon S. L. & Gaither S. E. (2018) Learning and Socializing Preferences in Hong Kong Chinese Children. *Child Development* **89**(6):2109–17. doi: 10.1111/cdev.13083. [NMcL]

Childers T. (2018) *The Third Reich: A history of Nazi Germany*. New York: Simon & Schuster. [DS]

Choi J. K. & Bowles S. (2007) The coevolution of parochial altruism and war. *Science* **318**(5850):636–40. Available at: http://science.sciencemag.org/content/318/5850/636.full. [aCKWDD, RH, NM]

Choi J. P., Chowdhury S. M. & Kim J. (2016) Group contests with internal conflict and power asymmetry. *Scandinavian Journal of Economics* **118**:816–40. [aCKWDD]

Chowdhury S. M., Jeon J. Y. & Ramalingam A. (2018) Property rights and loss aversion in contests. *Economic Inquiry* **56**(3):1492–511. [SMC, aCKWDD]

Chowdhury S. M., Lee D. & Sheremeta R. M. (2013) Top guns may not fire: Best-shot group contests with group-specific public good prizes. *Journal of Economic Behavior & Organization* **92**:94–103. [aCKWDD, RMS]

Chowdhury S. M. & Sheremeta R. M. (2011) A generalized Tullock contest. *Public Choice* **147**(3/4):413–20. [SMC, RMS]

Chowdhury S. M., Sheremeta R. M. & Turocy T. L. (2014) Overbidding and overspreading in rent-seeking experiments: Cost structure and prize allocation rules. *Games and Economic Behavior* **87**:224–38. [SMC]

Chowdhury S. M. & Topolyan I. (2016a) The attack-defense group contests: Best-shot versus weakest-link. *Economic Inquiry* **54**:548–57. [SMC, aCKWDD, RMS]

Chowdhury S. M. & Topolyan I. (2016b) Best-shot versus weakest-link in political lobbying: an application of group all-pay auction. *Social Choice and Welfare* **47**(4):959–71. [SMC]

Christensen C., Kern J. M., Bennitt E. & Radford A. N. (2016) Rival group scent induces changes in dwarf mongoose immediate behavior and subsequent movement. *Behavioral Ecology* **27**:1627–34. [ANR]

Christensen C. & Radford A. N. (2018) Dear enemies or nasty neighbours? Causes and consequences of variation in the responses of group-living species to territorial intrusions. *Behavioral Ecology* **29**:1004–13. [ANR, ARR]

Cikara M. & Van Bavel J. J. (2014) The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science* **9**:245–74. [aCKWDD]

Clark D. J. & Konrad K. A. (2007) Asymmetric conflict: Weakest link against best shot. *Journal of Conflict Resolution* **51**:457–69. [SMC, aCKWDD, RMS]

Clements K. & Schumacher J. A. (2010) Perceptual biases in social cognition as potential moderators of the relationship between alcohol and intimate partner violence: A review. *Aggression and Violent Behavior* **15**(5):357–68. [DK]

Coan J. A. & Allen J. J. (2003) Frontal EEG asymmetry and the behavioral activation and inhibition systems. *Psychophysiology* **40**:106–14. [aCKWDD]

Cohen-Chen S., Halperin E., Crisp R. J. & Gross J. J. (2014) Hope in the Middle East: Malleability beliefs, hope, and the willingness to compromise for peace. *Social Psychology and Personality Science* **5**:67–75. [rCKWDD]

Colman A. M. (1999) *Game theory and its applications in the social and biological sciences*. Routledge. [MWK]

Colman A. M. (2003) Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences* **26**:139–98. [aCKWDD]

Coombs C. H. (1973) A reparameterization of the prisoner's dilemma game. *Behavioral Science* **18**:424–28. [aCKWDD]

Coombs C. H. & Avrunin G. S. (1988) *The structure of conflict*. Lawrence Erlbaum. [aCKWDD]

Corr P. J. (ed.) (2008) *The reinforcement sensitivity theory of personality*. Cambridge University Press. [NMcN]

Corr P. J. (2016) Reinforcement Sensitivity Theory of Personality Questionnaires: Structural survey with recommendations. *Personality and Individual Differences* **89**:60–64. doi:10.1016/j.paid.2015.09.045. [NMcN]

Corr P. J., DeYoung C. G. & McNaughton N. (2013) Motivation and personality: A neuropsychological perspective. *Social and Personality Psychology Compass* **7**:158–75. [NMcN]

Corr P. J. & McNaughton N. (2012) Neuroscience and approach/avoidance personality traits: A two stage (valuation-motivation) approach. *Neuroscience and Biobehavioral Reviews* **36**:2339–54. doi:10.1016/j.neubiorev.2012.09.013. [NMcN]

Correll J. & Park B. (2005) A model of the in-group as a social resource. *Personality and Social Psychology Review* **9**:341–59. [aCKWDD]

Cosmides L. & Tooby J. (1987) From evolution to behaviour: Evolutionary psychology as the missing link. In: *The latest on the best: Essays on evolution and optimality*, ed. J. Dupre, pp. 276–306. MIT Press. [aCKWDD]

Costello K. & Hodson G. (2014) Explaining dehumanization among children: The interspecies model of prejudice. *British Journal of Social Psychology* **53**(1):175–97. doi: 10.1111/bjso.12016. [NMcL]

Cottrell C. A. & Neuberg S. L. (2005) Different emotional reactions to different groups: A sociofunctional threat-based approach to "prejudice." *Journal of Personality and Social Psychology* **88**(5):770–89. [DK]

Craig M. A. & Richeson J. A. (2014) More diverse yet less tolerant? How the increasingly diverse racial landscape affects white Americans' racial attitudes. *Personality and Social Psychology Bulletin* **40**:750–61. [PP]

Crawford V. P. & Iriberri N. (2007) Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games. *American Economic Review* **97**:1731–50. [rCKWDD]

Crofoot M. C. & Gilby I. C. (2012) Cheating monkeys undermine group strength in enemy territory. *Proceedings of the National Academy of Sciences USA* **109**(2):501–505. [TMA]

Crofoot M. C., Gilby I. C., Wikelski M. C. & Kays R. W. (2008) Interaction location outweighs the competitive advantage of numerical superiority in *Cebus capucinus* intergroup contests. *Proceedings of the National Academy of Sciences USA* **105**:577–81. [ANR]

Cunningham D. E., Gleditsch K. S. & Salehyan I. (2009) It takes two: A dyadic analysis of civil war duration and outcome. *Journal of Conflict Resolution* **53**:570–97. [aCKWDD]

Currarini S., Jackson M. O. & Pin P. (2009) An economic model of friendship: Homophily, minorities, and segregation. *Econometrica* **77**(4):1003–45. [DS]

Daly M. & Wilson M. (1988) *Homicide*. A. de Gruyter. [ACL]

Darwin C. (1859) On the origin of species by means of natural selection: Or: The 778 preservation of favoured races in the struggle for life. John Murray. [RMS]

Darwin C. (1873) *The descent of man*. Appleton. [aCKWDD]

Davidson R. J., Putnam K. M. & Larson C. L. (2000) Dysfunction in the neural circuitry of emotion regulation – A possible prelude to violence. *Science* **289**(5479):591–94. Available at: https://doi.org/10.1126/science.289.5479.591. [PCC]

Dawkins R. & Krebs J. R. (1979) Arms races between and within species. *Proceedings of the Royal Society B: Biological Sciences* **205**:489–511. [arCKWDD]

De Dreu C. K. W. (2010) Social conflict: The emergence and consequences of struggle and negotiation. In: *Handbook of social psychology*, ed. S. T. Fiske, D. T. Gilbert & H. Lindzey, 5th edition, vol. **2**, pp. 983–1023. Wiley. [aCKWDD]

De Dreu C. K. W. (2012) Oxytocin modulates cooperation within and competition between groups: An integrative review and research agenda. *Hormones and Behavior* **61**:419–28. [aCKWDD]

De Dreu C. K. W., Balliet D. & Halevy N. (2014) Parochial cooperation in humans: Forms and functions of self-sacrifice in intergroup competition and conflict. *Advances in Motivational Science* **1**:1–47. [aCKWDD]

De Dreu C. K. W. & Carnevale P. J. (2003) Motivational bases of information processing and strategy in conflict and negotiation. *Advances in Experimental Social Psychology* **35**:235–91. [aCKWDD]

De Dreu C. K. W., Carnevale P. J. D., Emans B. J. M. & Van de Vliert E. (1994) Effects of gain-loss frames in negotiation: Loss aversion, mismatching, and frame adoption. *Organizational Behavior and Human Decision Processes* **60**:90–107. [aCKWDD]

De Dreu C. K. W., Evers A., Beersma B., Kluwer E. S. & Nauta A. (2001) A theory-based measure of conflict management strategies in the workplace. *Journal of Organizational Behavior* **22**:645–68. [aCKWDD]

De Dreu C. K. W., Giacomantonio M., Giffin M. R. & Vecchiato G. (2019) Psychological constraints on aggressive predation in economic contests. *Journal of Experimental Psychology: General*. Available at: http://dx.doi.org/10.1037/xge0000531. [arCKWDD]

De Dreu C. K. W. & Giffin M. R. (2018) Hormonal modulation of attacker-defender contests. Unpublished manuscript, Leiden University. [arCKWDD]

De Dreu C. K. W., Greer L. L., Handgraaf M. J. J., Shalvi S., Van Kleef G. A., Baas M., Ten Velden F. S., Van Dijk E. & Feith S. W. W. (2010) The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* **328**:1408–11. [arCKWDD, RH]

De Dreu C. K. W., Greer L. L., Van Kleef G. A., Shalvi S. & Handgraaf M. J. (2011) Oxytocin promotes human ethnocentrism. *Proceedings of the National Academy of Sciences USA* **108**(4):1262–66. Available at: https://www.pnas.org/content/108/4/1262. [RH, rCKWDD]

De Dreu C. K. W., Gross J., Meder Z., Griffin M. R., Prochazkova E., Krikeb J. & Columbus S. (2016a) In-group defense, out-group aggression, and coordination failure in intergroup conflict. *Proceedings of the National Academy of Sciences USA* **113**:10524–29. [arCKWDD, DK, OW]

De Dreu C. K. W., Kluwer E. S. & Nauta A. (2008) The structure and management of conflict: Fighting or defending the status quo. *Group Processes and Intergroup Relations* **11**:331–53. [aCKWDD]

De Dreu C. K. W., Koole S. L. & Steinel W. (2000) Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation. *Journal of Personality and Social Psychology* **79**(6):975–87. [rCKWDD, NH]

De Dreu C. K. W. & Kret M. E. (2016) Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense. *Biological Psychiatry* **79** (3):165–73. Available at: https://doi.org/10.1016/j.biopsych.2015.03.020. [TOP]

De Dreu C. K. W., Kret M. E. & Sligte I. G. (2016b) Modulating prefrontal control in humans reveals distinct pathways to competitive success and collective waste. *Social Cognitive and Affective Neuroscience* **11**:1236–44. [arCKWDD]

De Dreu C. K. W. & McCusker C. (1997) Gain-loss frames and cooperation in two-person social dilemmas: A transformational analysis. *Journal of Personality and Social Psychology* **72**:1093–106. [aCKWDD]

De Dreu C. K. W., Nauta A. & Van de Vliert E. (1995) Self-serving evaluation of conflict behavior and escalation of the dispute. *Journal of Applied Social Psychology* **25**:2049–66. [aCKWDD]

De Dreu C. K. W., Scholte H. S., Van Winden F. A. A. M. & Ridderinkhof K. R. (2015) Oxytocin tempers calculated greed but not impulsive defense in predator-prey contests. *Social Cognitive and Affective Neuroscience* **5**:721–28. [aCKWDD]

De Dreu C. K. W., Weingart L. R. & Kwon S. (2000) Influence of social motives on integrative negotiation: A meta-analytical review and test of two theories. *Journal of Personality and Social Psychology* **78**:889–905. [aCKWDD]

De Juan A. (2015) Long-term environmental change and geographical patters of violence in Dafur, 2003–2005. *Political Geography* **45**:22–33. [aCKWDD]

De la Rosa L. E. (2011) Overconfidence and moral hazard. *Games and Economic Behavior* **73**:429–51. [aCKWDD]

Decety J. & Cowell J. M. (2014) The complex relation between morality and empathy. *Trends in Cognitive Sciences* **18**:337–9. [aCKWDD]

Dechenaux E., Kovenock D. & Sheremeta R. M. (2015) A survey of experimental research on contests, all-pay auctions, and tournaments. *Experimental Economics* **18**(4):609–69. [aCKWDD, MWK, RMS, SMC]

Deck C. & Sheremeta R. M. (2012) Fight or flight? Defending against sequential attacks in the game of siege. *Journal of Conflict Resolution* **56**(6):1069–88. [MWK]

Depue R. A. & Collins P. F. (1999) Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences* **22**:491–523. [aCKWDD]

Deutsch M. (1973) *The resolution of conflict*. Yale University Press. [aCKWDD]

Doğan G., Glowacki L. & Rusch H. (2018) Spoils division rules shape aggression between natural groups. *Nature Human Behaviour* **2**(5):322–26. Available at: https://doi.org/10.1038/s41562-018-0338-z. [aCKWDD, HR]

Donaldson Z. R. & Young L. J. (2008) Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* **322**(5903):900–904. Available at: http://science.sciencemag.org/content/322/5903/900.long. [RH]

Dore R. A., Hoffman K. M., Lillard A. S. & Trawalter S. (2014) Children's racial bias in perceptions of others' pain. *British Journal of Developmental Psychology* **32**(2):218–31. doi: 10.1111/bjdp.12038. [NMcL]

Dorris M. C. & Glimcher P. W. (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* **44**(2):365–78. [MWK]

Dosenbach N. U. F., Fair D. A., Cohen A. L., Schlaggar B. L. & Petersen S. E. (2008) A dual-networks architecture of top-down control. *Trends in Cognitive Sciences* **12**:99–105. [aCKWDD]

Dresher M. (1962) A sampling inspection problem in arms control agreements: A game-theoretic analysis. Memorandum RM-2972-ARPA. RAND Corp. [aCKWDD]

Drislane L. E. & Patrick C. J. (2017) Integrating alternative conceptions of psychopathic personality: A latent variable model of triarchic psychopathy constructs. *Journal of Personality Disorders* **31**(1):110–32. Available at: https://doi.org/10.1521/pedi_2016_30_240. [TOP]

Duffy J. & Kim M. (2005) Anarchy in the laboratory (and the role of the state). *Journal of Economic Behavior and Organization* **56**:297–329. [rCKWDD]

Dufwenberg M., Gächter S. & Hennig-Schmidt H. (2011) The framing of games and the psychology of play. *Games and Economic Behavior* **73**:459–78. [aCKWDD]

Dugatkin L. A. & Godin J. G. (1992) Prey approaching predators: A cost-benefit perspective. *Annals Zoologica Fennici* **29**:233–52. [aCKWDD]

Dunbar R. I. (1993) Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* **16**:681–94. [JO]

Dunham Y., Baron A. S. & Carey S. (2011) Consequences of "minimal" group affiliations in children. *Child Development* **82**(3):793–811. doi: 10.1111/j.1467-8624.2011.01577.x. [NMcL]

Durham W. H. (1976) Resource competition and human aggression: 1. Review of primitive war. *Quarterly Review of Biology* **51**:385–415. [aCKWDD]

Durham Y., Hirshleifer J. & Smith V. (1998) Do the rich get richer and the poor poorer? Experimental tests of a model of power. *American Economic Review* **88**:970–83. [aCKWDD]

Earle T. (1987) Chiefdoms in archaeological and ethno-historical perspective. *Annual Review of Anthropology* **16**:279–308. [aCKWDD]

Eckstein M., Becker B., Scheele D., Scholz C., Preckel K., Schlaepfer T. E., Grinevich V., Kendrick K. M., Maier W. & Hurlemann R. (2015) Oxytocin facilitates the extinction of conditioned fear in humans. *Biological Psychiatry* **78**(3):194–202. Available at: https://www.sciencedirect.com/science/article/pii/S0006322314007951?via%3Dihub. [RH]

Edens J. F., Marcus D. K., Lilienfeld S. O. & Poythress N. G. (2006) Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology* **115**(1):131–44. Available at: https://doi.org/10.1037/0021-843X.115.1.131. [TOP]

Efferson C., Lalive R. & Fehr E. (2008) The coevolution of cultural groups and in-group favoritism. *Science* **321**:1844–49. [aCKWDD]

Egas M. & Riedl A. (2008) The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences* **275**:871–78. [aCKWDD]

Eisenegger C., Haushofer J. & Fehr E. (2011) The role of testosterone in social interaction. *Trends in Cognitive Science* **15**:263–71. [aCKWDD]

Eliaz K. & Rubinstein A. (2011) Edgar Allan Poe's riddle: Framing effects in repeated matching pennies games. *Games and Economic Behavior* **71**(1):88–99. [rCKWDD, MWK]

Ellingsen T., Johannesson M., Tjotta S. & Torsvik G. (2010) Testing guilt aversion. *Games and Economic Behavior* **68**:95–107. [aCKWDD]

Elliot A. J. & Church M. A. (1997) A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology* **72**:218–32. [aCKWDD]

Engelmann D. & Strobel M. (2004) Inequality aversion, efficiency, and maximum preferences in simple distribution experiments. *American Economic Review* **94**:857–69. [aCKWDD]

Enquist M. & Leimar O. (1983) Evolution of fighting behaviour: decision rules and assessment of relative strength. *Journal of Theoretical Biology* **102**:387–410. [ANR]

Epps J. & Kendall P. C. (1995. Hostile attributional bias in adults. *Cognitive Therapy and Research* **19**(2):159–78. [DK]

Everett J. A., Pizarro D. A. & Crockett M. J. (2016) Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General* **145**(6):772–787. [AM]

Falk D. & Hildebolt C.F. (2017) Annual war deaths in small-scale versus state societies scale with population size rather than violence. *Current Anthropology* **58**(6):805–13. [WB, aCKWDD]

Fearon J. D., Humphreys M. & Weinstein J. M. (2009) Can development aid contribute to social cohesion after civil was? Evidence from a field experiment in post-conflict Liberia. *American Economic Review* **99**:287–91. [aCKWDD]

Fehr E. & Fischbacher U. (2004) Social norms and human cooperation. *Trends in Cognitive Sciences* **8**(4):185–90. Available at: https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(04)00050-6?_returnURL=https://%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1364661304000506%3Fshowall%3Dtrue. [RH]

Fehr E. & Gächter S. (2000) Cooperation and punishment in public goods experiments. *American Economic Review* **90**:980–94. [aCKWDD]

Fehr E. & Gächter S. (2002) Altruistic punishment in humans. *Nature* **415**(6868):137–40. [aCKWDD, RH]

Fehr E. & Schmidt K. M. (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* **114**:817–68. [aCKWDD, RMS]

Feldman S., Huddy L. & Marcus G. E. (2015) *Going to war in Iraq: When citizens and the press matter*. University of Chicago Press. [TMA]

Fiedler K. (2000) Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review* **107**:659–76. [aCKWDD]

Filiz-Ozbay E. & Ozbay E. Y. (2007) Auctions with anticipated regret: Theory and experiment. *American Economic Review* **97**:1407–18. [RMS]

Fischer R., Callander R., Reddish P. & Bulbulia J. (2013) How do rituals affect cooperation? An experimental field study comparing nine ritual types. *Human Nature* **24**(2):115–25. [aCKWDD]

Fiske A. P. (1992) The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review* **99**:689–723. [aCKWDD]

Fiske A. P. & Tetlock P. E. (1997) Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. *Political Psychology* **18**:255–97. [aCKWDD]

Fiske S. T. (1980) Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology* **38**:889–906. [aCKWDD]

Fjelde H. (2015) Farming or fighting? Agricultural price shocks and civil war in Africa. *World Development* **67**:525–34. [aCKWDD]

Flood M. M. (1972) The hide and seek game of Von Neumann. *Management Science* **18**:107–109. [aCKWDD]

Fog A. (2017) *Warlike and peaceful societies: The interaction of genes and culture*. Open Book. Available at: https://www.openbookpublishers.com/product/657/. [AF]

Fontaine J. R. J., Scherer K. R., Roesch E. B. & Ellsworth P. C. (2007) The world of emotions is not two-dimensional. *Psychological Science* **18**(12):1050–1057. https://doi.org/10.1111/j.1467-9280.2007.02024.x. [PCC]

Ford R. & Blegen M. (1992) Offensive and defensive use of punitive tactics in explicit bargaining. *Social Psychology Quarterly* **55**:351–62. [aCKWDD]

Fowles D. C. (2018) Temperament risk factors for psychopathy. In: *Handbook of psychopathy*, ed. C. J. Patrick, 2nd ed., pp. 94–126. New York: Guilford Press. [TOP]

Fowles D. C. & Dindo L. (2006) A dual-deficit model of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy*, ed. C. J. Patrick, 1st ed., pp. 14–34. Guilford Press. [TOP]

Frank R. (1988) *Passions within reason: The strategic role of the emotions*, vol. **1**. Norton. [AM]

Franke J., Kanzow C., Leininger W. & Schwartz A. (2013) Effort maximization in asymmetric contest games with heterogeneous contestants. *Economic Theory* **52**(2):589–630. [rCKWDD, MWK]

Freedman L. (2015) *Strategy: A history*. Oxford University Press. [DS]

Freire P. (1970) *Pedagogy of the oppressed*. Herder and Herder. [NS]

Frimer J., Motyl M. & Tell C. (2017) Sacralizing liberals and fair-minded conservatives: Ideological symmetry in the moral motives in the culture war. *Analyses of Social Issues and Public Policy* **17**(1):33–59. [AM]

Frost R. & McNaughton N. (2017) The neural basis of delay discounting: A review and preliminary model. *Neuroscience and Biobehavioral Reviews* **79**:48–65. doi:10.1016/j.neubiorev.2017.04.022. [NMcN]

Fu F., Tarnita C. E., Christakis N. A., Wang L., Rand D. G. & Nowak M. A. (2012) The evolution of in-group favoritism. *Scientific Reports* **2**:460. [aCKWDD]

Furrer R. D., Kyabulima S., Willems E. P., Cant M. A. & Manser M. B. (2011) Location and group size influence decisions in simulated intergroup encounters in banded mongooses. *Behavioral Ecology* **22**:493–500. [ANR]

Gächter S., Nosenzo D. & Sefton M. (2013) Peer effects in pro-social behavior: Social norms or social preferences? *Journal of the European Economic Association* **11**:548–73. [aCKWDD]

Gaither S. E., Chen E. E., Corriveau K. H., Harris P. L., Ambady N. & Sommers S. R. (2014) Monoracial and biracial children: effects of racial identity saliency on social learning and social preferences. *Child Development* **85**(6):2299–316. doi: 10.1111/cdev.12266. [NMcL]

Galanter N., Silva D., Rowell J. T. & Rychtářc J. (2017) Resource competition amid overlapping territories: The territorial raider model applied to multi-group interactions. *Journal of Theoretical Biology* **412**:100–106. [aCKWDD]

Gallup G. G. (1965) Aggression in rats as a function of frustrative nonreward in a straight alley. *Psychonomic Science* **3**:99–100. [NMcN]

Garcia J. & van den Bergh J. C. J. M. (2011) Evolution of parochial altruism by multilevel selection. *Evolution and Human Behavior* **32**:277–87. [aCKWDD]

Garcia J., Van Veelen M. & Traulsen A. (2014) Evil green beards: Tag recognition can also be used to withhold cooperation in structured populations. *Journal of Theoretical Biology* **360**:181–86. [aCKWDD]

Garfinkel M. R. & Skaperdas S. (2007) Economics of conflict: An overview. *Handbook of Defense Economics* **2**:649–709. [RMS]

Gat A. (1999) The pattern of fighting in simple, small-scale, prestate societies. *Journal of Anthropological Research* **55**(4):563–583. [WB]

Gat A. (2006) *War in human civlization*. Oxford University Press. Retrieved from http://www.loc.gov/catdir/toc/ecip0614/2006017223.html. [ACL]

Gavrilets S. & Fortunato L. (2014) A solution to the collective action problem in between-group conflict with within-group inequality. *Nature Communications* **5**:Article No. 3526. Available at: https://doi.org/10.1038/ncomms4526. [aCKWDD, AF]

Gavrilets S. & Richerson P. J. (2017) Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences USA* **114**(23):6068–73. Available at: https://www.pnas.org/content/early/2017/05/18/1703857114.full. [RH]

Geary D.C. (2000) Evolution and proximate expression of human paternal investment. *Psychological Bulletin* **126**:55–57. [DVB]

Gelfand M. J., LaFree G., Fahey S. & Feinberg E. (2013) Culture and extremism. *Journal of Social Issues* **69**:495–517. [aCKWDD]

Gelfand M. J., Raver J. L., Nishii L., Leslie L. M., Lun J., Lim B. C., Duan L., Almaliach A., Ang S., Arnadottir J., Aycan Z., Boehnke K., Boski P., Cabecinhas R., Chan D., Chhokar J., D'Amato A., Ferrer M., Fischlmayr I. C., Fischer R., Fülöp M., Georgas J., Kashima E. S., Kashima Y., Kim K., Lempereur A., Marquez P., Othman R., Overlaet B., Panagiotopoulou P., Peltzer K., Perez-Florizno LR, Ponomarenko L, Realo A, Schei V, Schmitt M, Smith P. B., Soomro N., Szabo E., Taveesin N., Toyama M., Van de Vliert E., Vohra N., Ward C. & Yamaguchi S. (2011) Differences between tight and loose cultures: A 33-nation study. *Science* **332**(6033):1100–104. Available at: https://doi.org/10.1126/science.1197754. [AF]

Gelfand M. J. & Realo A. (1999) Individualism-collectivism and accountability in intergroup negotiations. *Journal of Applied Psychology* **84**:721–736. [rCKWDD]

Gelpi C. (1997) Democratic diversions: Governmental structure and the externalization of domestic conflict. *Journal of Conflict Resolution* **41**(2):255–82. [TMA]

Gerber M. M. & Jackson J. (2013) Retribution as revenge and retribution as just deserts. *Social Justice Research* **26**(1):61–80. [DK]

Giancola P. R. (1995) Evidence for dorsolateral and orbital prefrontal cortical involvement in the expression of aggressive behavior. *Aggressive Behavior* **21**(6):431–50. Available at: https://doi.org/10.1002/1098-2337(1995)21:6<431::AID-AB2480210604>3.0.CO;2-Q. [PCC]

Giebels E., De Dreu C. K. W. & Van de Vliert E. (2000) Interdependence in negotiation: Impact of exit options and social motives on distributive and integrative negotiation. *European Journal of Social Psychology* **30**:255–72. [rCKWDD]

Gigerenzer G. & Brighton H. (2009) Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science* **1**:107–43. [aCKWDD]

Gigerenzer G. & Goldstein D. G. (1996) Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* **103**:650–69. [aCKWDD]

Gigerenzer G. & Selten R. (2002) *Bounded rationality: The adaptive toolbox*. MIT Press. [aCKWDD]

Gilby I. C., Brent L. J., Wroblewski E. E., Rudicell R. S., Hahn B. H., Goodall J. & Pusey A. E. (2013) Fitness benefits of coalitionary aggression in male chimpanzees. *Behavioral Ecology and Sociobiology* **67**(3):373–81. [WB]

Gilead M. & Lieberman N. (2014) We take care of our own: Caregiving salience increases out-group bias in response to out-group threat. *Psychological Science* **25**:1380–87. [aCKWDD]

Ginges J. & Atran S. (2011) War as a moral imperative (not just practical politics by other means). *Proceedings of the Royal Society B: Biological Sciences* **278**(1720):2930–38. Available at: https://doi.org/10.1098/rspb.2010.2384. [ACL]

Glowacki L., Isakov A., Wrangham R. W., McDermott R., Fowler J. H. & Christakis N. A. (2016) Formation of raiding parties for intergroup violence is mediated by social network structure. *Proceedings of the National Academy of Sciences USA* **113**:12114–19. [aCKWDD]

Glowacki L. & Von Rueden C. (2015) Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**(1683):20150010. [aCKWDD]

Glowacki L., Wilson M. & Wrangham R. (2017) The evolutionary anthropology of war. *Journal of Economic Behavior and Organization*. Advance online publication. Available at: https://doi.org/10.1016/j.jebo.2017.09.014. [HR]

Glowacki L. & Wrangham R.W. (2013) The role of rewards in motivating participation in simple warfare. *Human Nature* **24**(4):444–60. Available at: https://doi.org/10.1007/s12110-013-9178-8. [TMA, WB, ACL]

Glowacki L. & Wrangham R.W. (2015) Warfare and reproductive success in a tribal population. *Proceedings of the National Academy of Sciences USA* **112**(2):348–353. [WB]

Gneezy A. & Fessler D. M. (2012) Conflict, sticks and carrots: War increases prosocial punishments and rewards. *Proceedings of the Royal Society B: Biological Sciences* **279**:219–23. [aCKWDD]

Gochman C. S. & Maoz Z. M. (1984) Militarized interstate disputes, 1816–1976: Procedures, patterns, and insights. *Journal of Conflict Resolution* **28**:585–616. [aCKWDD]

Goddard S. E. (2015) The rhetoric of appeasement: Hitler's legitimation and British foreign policy, 1938–39. *Security Studies* **24**:95–130. [aCKWDD]

Goeree J. K., Holt C. A. & Palfrey T. R. (2003) Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior* **45**:97–113. [arCKWDD, MWK]

Goetz J., Keltner D. & Simon-Thomas E. (2010) Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin* **136**(3):351–74. Available at: https://doi.org/10.1037/a0018807. [PCC]

Goldstein J. S. (2003) *War and gender: How gender shapes the war system and vice versa*. Cambridge University Press. [ACL]

Gómez Á., López-Rodríguez L., Sheikh H., Ginges J., Wilson L., Waziri H., Vázquez A., Davis R. & Atran S. (2017) The devoted actor's will to fight and the spiritual dimension of human conflict. *Nature Human Behaviour* **1**(9):673–79. Available at: https://doi.org/10.1038/s41562-017-0193-3. [AM]

Goodwin G. P. & Darley J. M. (2010) The perceived objectivity of ethical beliefs: Psychological findings and implications for public policy. *Review of Philosophy and Psychology* **1**(2):161–88. Available at: https://doi.org/10.1007/s13164-009-0013-4. [AM]

Goodwin G. P. & Darley J. M. (2012) Why are some moral beliefs perceived to be more objective than others? *Journal of Experimental Social Psychology* **48**(1):250–56. Available at: https://doi.org/10.1016/j.jesp.2011.08.006. [AM]

Gould R. V. (1999) Collective violence and group solidarity: Evidence from a feuding society. *American Sociological Review* **64**:356–80. [aCKWDD]

Gould R. V. (2000) Revenge as sanction and solidarity display: An analysis of vendettas in nineteenth-century Corsica. *American Sociological Review* **65**:682–704. [aCKWDD]

Graham J., Haidt J., Koleva S. P., Motyl M., Iyer R., Wojcik S. P. & Ditto P H. (2013) Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology* **47**:55–130. [AM]

Gray J. A. (1977) Drug effects on fear and frustration: Possible limbic site of action of minor tranquilizers. In: *Handbook of psychopharmacology: Vol. 8: Drugs, neurotransmitters and behaviour*, ed. L. L. Iversen, S. D. Iversen & S. H. Snyder, pp. 433–529. Plenum Press. [NMcN]

Gray J. A. (1990) Brain systems that mediate both emotion and cognition. *Cognition and Emotion* **4**:269–88. [aCKWDD]

Gray J. A. & McNaughton N. (2000) *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*, 2nd edition. Oxford University Press. [NMcN]

Green M. J. & Phillips M. L. (2004) Social threat perception and the evolution of paranoia. *Neuroscience and Biobehavioral Reviews* **28**:333–42. [aCKWDD]

Griesser M. & Ekman J. (2005) Nepotistic mobbing behavior in the Siberian jay, *Perisoreus infaustus*. *Animal Behavior* **69**:345–52. [aCKWDD]

Griskevicius V., Cantú S. M. & Vugt M. V. (2012) The evolutionary bases for sustainable behavior: Implications for marketing, policy, and social entrepreneurship. *Journal of Public Policy & Marketing* **31**:115–28. [JO]

Gross J. & De Dreu C. K. W. (2019a) Individual solutions to shared problems create a modern tragedy of the commons. *Science Advances* **5**(4):eaau7296. [rCKWDD]

Gross J. & De Dreu C. K. W. (2019b) The rise and fall of cooperation through reputation and group polarization. *Nature Communications* **10**:1–10. http://doi.org/10.1038/s41467-019-08727-8 [aCKWDD]

Gross J., Emmerling F., Vostroknutov A. & Sack A. T. (2018) Manipulation of prosociality and rule-following with non-invasive brain stimulation. *Scientific Reports* **8**(1):1–10. https://doi.org/10.1038/s41598-018-19997-5 [aCKWDD]

Gross J., Meder Z. Z., Okamoto-Barth S. & Riedl A. (2016) Building the Leviathan: Voluntary centralisation of punishment power sustains cooperation in humans. *Nature Scientific Reports* **6**:20767. [aCKWDD]

Gross J., Woelbert E., Zimmermann J., Okamoto-Barth S., Riedl A. & Goebel R. (2014) Value signals in the prefrontal cortex predict individual preferences across reward categories. *Journal of Neuroscience* **34**:7580–86. [aCKWDD]

Gross J. J. (1998) The emerging field of emotion regulation: An integrative review. *Review of General Psychology* **2**(3):271–99. Available at: https://doi.org/10.1037/1089-2680.2.3.271. [PCC]

Gross J. J. (2001) Emotion regulation in adulthood: Timing is everything. *Current Directions in Psychological Science* **10**(6):214–19. Available at: https://doi.org/10.1111/1467-8721.00152. [PCC]

Grossman H. I. & Kim M. (1996) Predation and production. In: *The political economy of conflict and appropriation*, ed. M. R. Garfinkel & S. Skaperdas, pp. 57–71. Cambridge University Press. [aCKWDD]

Grossman H. I. & Kim M. (2002) Predation and accumulation. *Journal of Economic Growth* **158**:393–407. [arCKWDD]

Guastella AJ, Mitchell PB, Dadds MR. (2008) Oxytocin increases gaze to the eye region of human faces. *Biological Psychiatry* **63**(1):3–5. Available at: https://www.biologicalpsychiatryjournal.com/article/S0006-3223(07)00617-8/fulltext. [RH]

Guay J., Ruscio J., Knight A. R. & Hare R. D. (2007) A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality 116(4):701–16. Available at: https://doi.org/10.1037/0021-843X.116.4.701. [TOP]

Guimond S., Crisp R. J., De Oliveira P., Kamiejski R., Kteily N., Kuepper B., Lalonde R. N., Levin S., Pratto F., Tougas F., Sidanius J. & Zick A. (2013) Diversity policy, social dominance, and intergroup relations: Predicting prejudice in changing social and political contexts. *Journal of Personality and Social Psychology* **104**:941–58. Available at: https://doi.org/10.1037/a0032069. [KU]

Gürerk O., Irlenbusch B. & Rockenbach B. (2006) The competitive advantage of sanctioning institutions. *Science* **312**:108–11. [aCKWDD]

Habash G. (2011) When kindness attacks: A Q&A with Barbara Oakley. Available at: https://www.publishersweekly.com/pw/by-topic/industry-news/tip-sheet/article/49893-when-kindness-attacks-a-q-a-with-barbara-oakley.html. [RH]

Haberli M. A., Aeschlimann P. B. & Milinski M. (2005) Sticklebacks benefit from closer predator inspection: An experimental test of risk assessment. *Ethology Ecology, and Evolution* **17**:249–59. [aCKWDD]

Hafer C. (2006) On the origins of property rights: Conflict and production in the state of nature. *The Review of Economic Studies* **73**:119–43. [CH]

Haidt J. (2012) *The righteous mind: Why good people are divided by politics and religion*. Pantheon/Random House. [AM, PP]

Halevy N. (2008) Team negotiation: Social, epistemic, economic, and psychological consequences of subgroup conflict. *Personality and Social Psychology Bulletin* **34**(12):1687–702. [NH]

Halevy N. (2016) Preemptive strikes: Fear, hope, and defensive aggression. *Journal of Personality and Social Psychology* **112**:224–37. [aCKWDD]

Halevy N. (2017) Preemptive strikes: Fear, hope, and defensive aggression. *Journal of Personality and Social Psychology* **112**(2):224–37. Available at: https://doi.org/10.1037/pspi0000077. [HR]

Halevy N., Bornstein G. & Sagiv L. (2008) "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science* **19**(4):405–11. [aCKWDD, NM]

Halevy N. & Chou E. Y. (2014) How decisions happen: Focal points and blind spots in interdependent decision making. *Journal of Personality and Social Psychology* **106**:398–417. [aCKWDD]

Halevy N., Chou E. Y., Cohen T. R. & Bornstein G. (2010) Relative deprivation and intergroup competition. *Group Processes & Intergroup Relations* **13**:685–700. [aCKWDD]

Halevy N., Chou E. Y. & Murnighan J. K. (2011) Games groups play: Mental models in intergroup conflict and negotiation. In: *Negotiation and groups*, eds. E. A. Mannix, M. A. Neale & J. R. Overbeck, pp. 79–107. Emerald Group. [NH, rCKWDD]

Halevy N., Chou E. Y. & Murnighan J. K. (2012) Mind games: the mental representation of conflict. *Journal of Personality and Social Psychology* **102**(1):132–48. [NH]

Halevy N. & Halali E. (2015) Selfish third parties act as peacemakers by transforming conflicts and promoting cooperation. *Proceedings of the National Academy of Sciences USA* **112**:6937–42. [aCKWDD]

Halevy N., Krebs T. & De Dreu C. K. W. (2019) Psychological situations illuminate the meaning of human behavior: Recent advances and application to social influence processes. *Social and Personality Psychology Compass* **13**(2):12437. [rCKWDD]

Halevy N., Sagiv L., Roccas S. & Bornstein G. (2006) Perceiving intergroup conflict: From game models to mental templates. *Personality and Social Psychology Bulletin* **32**(12):1674–89. [rCKWDD, NH]

Hall P. J., Chong W., McNaughton N. & Corr P. J. (2011) An economic perspective on the reinforcement sensitivity theory of personality. *Personality and Individual Differences* **51**:242–47. [NMcN]

Halperin E. (2014) Emotion, emotion regulation, and conflict resolution. *Emotion Review*, **6**(1):68–76. Available at: https://doi.org/10.1177/1754073913491844. [PCC]

Halperin E. (2016) *Emotions in conflict: Inhibitors and facilitators of peace making*. Routledge. [PCC]

Halperin E., Porat R., Tamir M. & Gross J. J. (2013) Can emotion regulation change political attitudes in intractable conflicts? From the laboratory to the field. *Psychological Science* **24**(1):106–11. Available at: https://doi.org/10.1177/095679761245257. [PCC]

Halperin E., Russell A. G., Trzesniewski K. H., Gross J. J. & Dweck C. S. (2011) Promoting the middle east peace process by changing beliefs about group malleability. *Science* **333**(6050):1767–69. Available at: https://doi.org/10.1126/science.1202925. [PCC, rCKWDD]

Hamilton W. D. (1964a) The genetical evolution of social behavior: I. *Journal of Theoretical Biology* **7**:1–16. [JO]

Hamilton W. D. (1964b) The genetical evolution of social behavior: II. *Journal of Theoretical Biology* **7**:17–52. [JO]

Hamilton W. D. (1971) Geometry for selfish herd. *Journal of Theoretical Biology* **31**:295–301. [aCKWDD]

Hampton A. N., Bossaerts P. & O'Doherty J. P. (2008) Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences USA* **105**(18):6741–46. [MWK]

Hardy I. C. W. & Briffa M. (2013) *Animal contests*. Cambridge University Press. [ANR]

Hare R. D. & Neumann C. S. (2008) Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology* **4**:217–46. [aCKWDD]

Harinck F., De Dreu C. K. W. & Van Vianen A. E. M. (2000) The impact of conflict issues on fixed-pie perceptions, problem solving, and integrative outcomes in negotiation. *Organizational Behavior and Human Decision Processes* **81**:329–58. [rCKWDD]

Harinck F. & Druckman D. (2017) Do negotiation interventions matter? Resolving conflicting interests and values. *Journal of Conflict Resolution* **61**:29–55. [rCKWDD]

Harmon-Jones E. & Allen J. J. (1998) Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology* **74**(5):1310–16. [TMA]

Harmon-Jones E. & Sigelman J. (2001) State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology* **80**(5):797–803. [TMA, aCKWDD]

Harris P. L. & Corriveau K. H. (2011) Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366** (1567):1179–87. doi: 10.1098/rstb.2010.0321. [NMcL]

Harris P. L., Koenig M. A., Corriveau K. H. & Jaswal V. K. (2018) Cognitive Foundations of Learning from Testimony. *Annual Review of Psychology* **69**(1):251–73. doi: 10.1146/annurev-psych-122216-011710. [NMcL]

Hasan-Aslih S., Netzer L., van Zomeren M., Saguy T., Tamir M. & Halperin E. (2018) When we want them to fear us: The motivation to influence outgroup emotions in collective action. *Group Processes & Intergroup Relations*. Available at: https://doi.org/10.1177/1368430218769744. [DK]

Haselton M. G. & Nettle D. (2006) The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review* **10**:47–66. [aCKWDD]

Haselton M. G., Nettle D. & Andrews P. W. (2015) The evolution of cognitive bias. In: *The handbook of evolutionary psychology*, ed. D. M. Buss, pp. 724–46. Wiley. Available at: https://doi.org/10.1002/9780470939376.ch25. [AM]

Haslam N. (2006) Dehumanization: An integrative review. *Personality and Social Psychology Review* **10**:252–64. [aCKWDD]

Hässler T., Shnabel N., Ullrich J., Arditti-Vogel A. & SimanTov-Nachlieli I. (2018) Individual differences in system justification predict power and morality-related needs in advantaged and disadvantaged groups in response to group disparity. *Group Processes and Intergroup Relations*. doi:10.1177/1368430218773403. [NS]

Henrich J. & McElreath R. (2003) The evolution of cultural evolution. *Evolutionary Anthropology* **12**:123–35. [aCKWDD]

Henrich J., McElreath R., Barr A., Ensminger J., Barrett C., Bolyanatz A., Cardenas J. C., Gurven M., Gwako E., Henrich N., Lesorogol C., Marlowe F. W., Tracer D. & Ziker J. (2006) Costly punishment across human societies. *Science* **312**:1767–70. [aCKWDD]

Heinrichs M., Baumgartner T., Kirschbaum C. & Ehlert U. (2003) Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biological Psychiatry* **54**:1389–98. Available at: https://www.sciencedirect.com/science/article/pii/S0006322303004657. [RH]

Helms J. E. (1990) *Black and white racial identity: Theory, research, and practice*. Greenwood Press. [NS]

Herek G. M. & McLemore K. A. (2013) Sexual prejudice. *Annual Review of Psychology* **64**:309–33. [NS]

Hermalin B. E. (1998) Toward an economic theory of leadership: Leading by example. *American Economic Review* **88**:1188–206. [aCKWDD]

Herrmann B., Thöni C. & Gächter S. (2008) Antisocial punishment across societies. *Science* **319**(5868):1362–67. [NM]

Herrmann R. K., Tetlock P. E. & Visser P. S. (1999) Mass public decisions on go to war: A cognitive-interactionist framework. *American Political Science Review* **93**(3):553–73. [TMA]

Heyman J. & Ariely D. (2004) Effort for payment: A tale of two markets. *Psychological Science* **15**:787–93. [aCKWDD]

Hibbing J., Smith K. & Alford J. (2015) Liberals and conservatives: Non-convertible currencies. *Behavioral and Brain Sciences* **38**:E145. [PP]

Higgins E. (1997) Beyond pleasure and pain. *American Psychologist* **52**:1280–300. [aCKWDD]

Higgins E. (2000) Making a good decision: Value from fit. *American Psychologist* **55**:1217–30. [aCKWDD]

Hill K. R., Walker R. S., Božičević M., Eder J., Headland T., Hewlett B., Hurtado A. M., Marlowe F., Wiessner P. & Wood B. (2011) Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science* **331**:1286–89. [JO]

Hirshleifer J. (1988) The analytics of continuing conflict. *Synthese* **76**:201–33. [aCKWDD]

Hirshleifer J. (1991) The paradox of power. *Economics & Politics* **3**:177–200. [arCKWDD]

Holbrooke R. (1999) *To end a war*. Modern Library. [rCKWDD]

Holt C.A., Kydd A., Razzolini L. & Sheremeta R.M. (2016) The paradox of misaligned profiling: Theory and experimental evidence. *Journal of Conflict Resolution* **60**:482–500. [RMS]

Houston A. I. & McNamara J. M. (2005) John Maynard Smith and the importance of consistency in evolutionary game theory. *Biology & Philosophy* **20**:933–50. [ANR]

Huang Y., Kendrick K. M., Zheng H. & Yu R. (2015) Oxytocin enhances implicit social conformity to both in-group and out-group opinions. *Psychoneuroendocrinology* **60**:114–19. Available at: https://www.sciencedirect.com/science/article/pii/S0306453015002073?via%3Dihub. [RH]

Huddy L., Mason L. & Aarøe L. (2015) Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review* **109**(1):1–17. [TMA]

Hüffmeier J., Zerres A., Freund P. A., Backhaus K., Trötschel R. & Hertel G. (2018) Strong or weak synergy? Revising the assumption of team-related advantages in integrative negotiations. *Journal of Management*. Advance online publication. doi:10.1177/0149206318770245, [JH]

Hughes A. L. (1988) *Evolution and human kinship*. Oxford University Press. [JO]

Humphreys M. & Weinstein J. N. (2006) Handling and manhandling civilians in civil war. *American Political Science Review* **100**:429–47. [aCKWDD]

Humphreys M. & Weinstein J. M. (2008) Who fights? The determinants of participation in civil war. *American Journal of Political Science* **52**:436–55. [aCKWDD]

Huo Y. J., Smith H. J., Tyler T. R. & Lind E. A. (1996) Identification and justice concerns: Is separatism the problem; is assimilation the answer? *Psychological Science* **7**:40–45. Available at: https://doi.org/10.1111/j.1467-9280.1996.tb00664.x. [KU]

Hurlemann R. (2017) Oxytocin-augmented psychotherapy: beware of context. *Neuropsychopharmacology* **42**(1):377. Available at: https://www.nature.com/articles/npp2016188. [RH]

Hurlemann R., Patin A., Onur O. A., Cohen M.X., Baumgartner T., Metzler S., Dziobek I., Gallinat J., Wagner M., Maier W. & Kendrick K. M. (2010) Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience* **30**(14):4999–5007. Available at: https://www.ncbi.nlm.nih.gov/m/pubmed/20371820/?i=4&from=hurlemann%202010. [RH]

Hurlemann R. & Marsh N. (2016) New insights into the neuroscience of human altruism. *Der Nervenarzt* **87**(11):1131–35. Available at: https://link.springer.com/article/10.1007%2Fs00115-016-0229-3. [RH]

Hurlemann R. & Marsh N. (2017) Deciphering the modulatory role of oxytocin in human altruism. *Reviews in the Neuroscience* **28**(4):335–42. Available at: https://www.degruyter.com/view/j/revneuro.2017.28.issue-4/revneuro-2016-0061/revneuro-2016-0061.xml. [RH]

Hurlemann R. & Scheele D. (2016) Dissecting the role of oxytocin in the formation and loss of social relationships. *Biological Psychiatry* **79**(3):185–93. Available at: https://www.sciencedirect.com/science/article/pii/S0006322315004369. [RH]

Huth P. & Russett B. (1984) What makes deterrence work: Cases from 1900 to 1980. *World Politics* **36**:496–526. [aCKWDD]

Ifcher J. & Zarghamee H. (2014) Affect and overconfidence: A laboratory investigation. *Journal of Neuroscience, Psychology and Economics* **7**:125–50. [aCKWDD]

Inglehart R., Moaddel M. & Tessler M. (2006) Xenophobia and in-group solidarity in Iraq: A natural experiment on the impact of insecurity. *Perspectives on Politics* **4** (3):495–505. Available at: https://doi.org/10.1017/S1537592706060324. [AF]

Insel T. R. (1997) A neurobiological basis of social attachment. *The American Journal of Psychiatry* **154**(6):726–35. Available at: https://ajp.psychiatryonline.org/doi/abs/10.1176/ajp.154.6.726. [RH]

Insel T.R. & Young L.J. (2001) The neurobiology of attachment. *Nature Reviews Neuroscience* **2**(2):129–36. Available at: https://www.nature.com/articles/35053579. [RH]

Insko C. A., Schopler J., Drigotas S. M., Graetz K. A., Kennedy J., Cox C. & Bornstein G. (1993) The role of communication in interindividual-intergroup discontinuity. *Journal of Conflict Resolution* **37**:108–38. doi:10.1177/0022002793037001005. [JH]

Insko C. A., Schopler J., Gaertner L., Wildschut T., Kozar R., Pinter B., Finkel E. J., Brazil D. M., Cecil C. L. & Montoya M. R. (2001) Interindividual–intergroup discontinuity reduction through the anticipation of future interaction. *Journal of Personality and Social Psychology* **80**:95–111. doi:10.1037//0022-3514.80.1.95. [JH]

Insko C. A., Schopler J., Hoyle R. H., Dardis G. J. & Graetz K. A. (1990) Individual-group discontinuity as a function of fear and greed. *Journal of Personality and Social Psychology* **58**:68–79. doi:10.1037/0022-3514.58.1.68. [JH]

Israel S., Weisel O., Ebstein R. P., Bornstein G. (2012) Oxytocin, but not vasopressin, increases both parochial and universal altruism. *Psychoneuroendocrinology* **37**:1341–44. Available at: https://www.sciencedirect.com/science/article/pii/S0306453012000546. [RH]

Izuma K., Adolphs R. (2013) Social manipulation of preference in the human brain. *Neuron* **78**(3):563–73. https://www.sciencedirect.com/science/article/pii/S0896627313002705?via%3Dihub. [RH]

Jackson J. C., Choi V. K. & Gelfand M. J. (2019) Revenge: A multilevel review and synthesis. *Annual Review of Psychology* **70**:319–45. [DK]

Jackson J. C., Jong J., Bilkey D., Whitehouse H., Zollmann S. & McNaughton C. & Halberstadt J. (2018) Synchrony and physiological arousal increase cohesion and cooperation in large naturalistic groups. *Scientific Reports* **8**:127. [aCKWDD]

Janis I. L. (1972) *Victims of groupthink: A psychological study of foreign-policy decisions and fiascos*. Houghton Mifflin. [aCKWDD]

Janssen M., Anderies J. M. & Joshi S. R. (2011) Coordination and cooperation in asymmetric commons dilemmas. *Experimental Economics* **14**:547–66. [aCKWDD]

Jentleson B. W. (1992) The pretty prudent public: Post post-Vietnam American opinion on the use of military force. *International Studies Quarterly* **36**(1):49–74. [TMA]

Jentleson B. W. & Britton R. L. (1998) Still pretty prudent: Post-Cold War American public opinion on the use of military force. *Journal of Conflict Resolution* **42**(4):395–417. [TMA]

Jervis R. (1978) Cooperation under the security dilemma. *World Politics* **30**:167–214. [aCKWDD]

Johnson D. D. P. (2004) *Overconfidence and war: The havoc and glory of positive illusions*. Harvard University Press. [aCKWDD]

Johnson D. D. P. (2005) God's punishment and public good – A test of the supernatural punishment hypothesis in 186 world cultures. *Human Nature* **16**:410–46. [aCKWDD]

Johnson D. D. P. (2006) Overconfidence in war games: Experimental evidence on expectations, aggression, gender, and testosterone. *Proceedings of the Royal Society B: Biological Sciences* **273**:2513–20. [aCKWDD]

Johnson D. D. P. & Fowler J. H. (2011) The evolution of overconfidence. *Nature* **477**:316–20. [aCKWDD]

Johnstone R. A. (2001) Eavesdropping and animal conflict. *Proceedings of the National Academy of Sciences USA* **98**:9177–80. [ANR]

Jones D. (2018) Kin selection and ethnic group selection. *Evolution and Human Behavior* **39**(1):9–18. [DS]

Jones D. M., Bremer S. A. & Singer J. D. (1996) Militarized interstate disputes 1816–1992: Rationale, coding rules, and empirical patterns. *Conflict Management and Peace Science* **15**:163–215. [aCKWDD]

Jost J. T. & Banaji M. R. (1994) The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology* **33**:1–27. [NS]

Kagel J. H. & Roth A. E. (1995) *Handbook of experimental economics*. Princeton University Press. [aCKWDD]

Kahneman D. & Klein G. (2009) Conditions for intuitive expertise: A failure to disagree. *American Psychologist* **64**:515–26. [aCKWDD]

Kahneman D., Knetsch J. L. & Thaler R. H. (1991) The endowment effect, loss aversion, and the status quo bias. *Journal of Economic Perspectives* **5**:193–206. [aCKWDD]

Kahneman D. & Renshon J. (2007) Why hawks win. *Foreign Policy* **158**:34–38. Available at: https://foreignpolicy.com/2009/10/13/why-hawks-win/. [NH]

Kahneman D. & Tversky A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica* **47**:263–91. [aCKWDD, NMcN]

Kahneman D. & Tversky A. (1984) Choices, values, and frames. *American Psychologist* **39**(4):341–50. [aCKWDD]

Kahneman D. & Tversky A. (1995) Conflict resolution: A cognitive perspective. In: *Barriers to conflict resolution*, ed. K. Arrow, R. H. Mnookin, L. Ross, A. Tversky & R. Wilson, pp. 44–61. Norton. [aCKWDD]

Kanai R., Feilden T., Firth C. & Rees G. (2011) Political orientations are correlated with brain structure in young adults. *Current Biology* **21**:677–80. [PP]

Kanat M., Heinrichs M., Mader I., van Elst L. T. & Domes G. (2015) Oxytocin modulates amygdala reactivity to masked fearful eyes. *Neuropsychopharmacology* **40**(11):2632–38. Available at: https://www.nature.com/articles/npp2015111. [RH]

Kelley H. H., Holmes J. G., Kerr N. L., Reis H. T., Rusbult C. E. & Van Lange P. A. M. (2003) *An atlas of interpersonal relations*. Cambridge University Press. [aCKWDD]

Kelley H. H. & Thibaut J. W. (1978) *Interpersonal relations: A theory of interdependence*. Wiley. [aCKWDD]

Kelly J. F. & Hake D. F. (1970) An extinction-induced increase in an aggressive response with humans. *Journal of the Experimental Analysis of Behavior* **14**:153–64. [NMcN]

Kelman H. (2006) Interests, relationships, identities: Three central issues for individuals and groups in negotiating their social environment. *Annual Review of Psychology* **57**:1–26. [rCKWDD]

Keltner D., Gruenfeld D. H. & Anderson C. (2003) Power, approach, and inhibition. *Psychological Review* **110**:265–84. [NS]

Keltner D. & Robinson R. J. (1997) Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin* **23**:1066–77. [aCKWDD]

Keynes J. M. (1919) *The economic consequences of the peace*. MacMillan. [aCKWDD]

Kimbrough E. O., Laughren K. & Sheremeta R. (2019) War and conflict in economics: Theories, applications, and recent trends. *Journal of Economic Behavior & Organization*. Available at: https://doi.org/10.1016/j.jebo.2017.07.026. [RMS]

Kinzler K. D., Dupoux E. & Spelke E. S. (2007) The native language of social cognition. *Proceedings of the National Academy of Sciences USA* **104**(30):12577–80. doi: 10.1073/pnas.0705345104 [NMcL]

Kitchen D. M. & Beehner J. C. (2007) Factors affecting individual participation in group-level aggression among non-human primates. *Behaviour* **144**(12):1551–81. [TMA, ANR]

Klandermans P. G. & van Stekelenburg J. (2013) Social movements and the dynamics of collective action. In: *The Oxford handbook of political psychology*, ed. L. Huddy, D. O. Sears & J. S. Levy, 2nd edition, pp. 774–812. [TMA]

Klavina L. & Buunk B. (2013) Intergroup intrasexual competition: Reactions towards outgroup members as romantic rivals. *Journal of Evolutionary Psychology* **11**:93–120. [DVB]

Klimecki O. (in press) The role of empathy in conflict resolution. *Emotion Review*. [PCC]

Klimecki O. M., Sander D. & Vuilleumier P. (2018) Distinct brain areas involved in anger versus punishment during social interactions. *Scientific Reports* **8**:10556. Available at: https://doi.org/10.1038/s41598-018-28863-3. [PCC]

Kluwer E. S., Heesink J. A. M. & Van de Vliert E. (1997) The marital dynamics of conflict over the division of labor. *Journal of Marriage and Family* **59**:635–53. [aCKWDD]

Knoch D., Gianotti L. R., Pascual-Leone A., Treyer V., Regard M., Hohmann M. & Brugger P. (2006a) Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *Journal of Neuroscience* **26**:6469–72. [aCKWDD]

Knoch D., Pascual L. A., Meyer K., Trever V. & Fehr E. (2006b) Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* **314**:829–32. [aCKWDD]

Koellinger P. & Treffers T. (2015) Joy leads to overconfidence, and a simple countermeasure. *PLoS One* **12**:e0143263. [aCKWDD]

Konrad K. A. (2009) *Strategy and dynamics in contests. Oxford University Press*. [RMS]

Konrad K. A. & Morath F. (2012) Evolutionary stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology* **306**:61–67. [aCKWDD]

Kosfeld M., Heinrichs M., Zak P. J., Fischbacher U. & Fehr E. (2005) Oxytocin increases trust in humans. *Nature* **435**(7042):673–76. Available at: https://www.nature.com/articles/nature03701. [RH]

Kovenock D. & Roberson B. (2012) Conflicts with multiple battlefields. In: *The Oxford handbook of the economics of peace and conflict*, ed. M. R. Garfinkel & S. Skaperdas, pp. 503–31. Oxford University Press. [SMC]

Kovenock D. & Roberson B. (2018) The optimal defense of networks of targets. *Economic Inquiry* **56**(4):2195–211. [RMS, SMC]

Kovenock D., Roberson B. & Sheremeta R. M. (2010) The attack and defense of weakest-link networks. CESifo Working Paper No. 3211. [MWK]

Kovenock D., Roberson B. & Sheremeta R. M. (2019) The attack and defense of weakest-link networks. *Public Choice* **179**(3/4):175–94. Available at: https://doi.org/10.1007/s11127-018-0618-1. [SMC, RMS]

Kramer R. M. (1995) Power, paranoia and distrust in organizations: The distorted view from the top. *Research on Negotiation in Organizations* **5**:119–54. [aCKWDD]

Kraus M. W., Piff P. K., Mendoza-Denton R., Rheinschmidt M. L. & Keltner D. (2012) Social class, solipsism, and contextualism: How the rich are different from the poor. *Psychological Review* **119**:546–72. [NS]

Krause J. & Ruxton G.D. (2002) *Living in groups*. Oxford University Press. [ARR]

Krol M. & Krol M. (2017) A novel approach to studying strategic decisions with eye-tracking and machine learning. *Judgment and Decision Making* **12**(6):596–609. [MWK]

Kteily N., Saguy T., Sidanus J. & Taylor D. M. (2013) Negotiating power: Agenda ordering and the willingness to negotiate in asymmetric intergroup conflicts. *Journal of Personality and Social Psychology* **105**(6):978–95. [aCKWDD, NH]

Kuhberger A. (1998) The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes* **75**:23–55. [aCKWDD]

Kull S., Ramsay C. & Lewis E. (2003) *Misperceptions, the media, and the Iraq war*. Political Science Quarterly *118(4):569–98. [TMA]*

Kydd A. (2011) Terrorism and profiling. *Terrorism and Political Violence* **23**:458–73. [RMS]

Lacomba J., Lagos F., Reuben E. & van Winden F. (2014) On the escalation and de-escalation of conflict. *Games and Economic Behavior* **86**:40–57. [arCKWDD]

Lamm C., Decety J. & Singer T. (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* **54**:2492–502. [aCKWDD]

Lang M., Bahma V., Shaver J. H., Reddish P. & Xygalatas D. (2017) Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology* **127**:191–97. [aCKWDD]

Langergraber K.E., Watts D.P., Vigilant L. & Mitani J.C. (2017) Group augmentation, collective action, and territorial boundary patrols by male chimpanzees. *Proceedings of the National Academy of Sciences USA* **114**(28):7337–42. [WB]

Lax D. A. & Sebenius J. K. (1986) *The manager as negotiator: Bargaining for cooperation and competitive gain*. Free Press. [rCKWDD]

LeBlanc S.A. (2016) Forager warfare and our evolutionary past. In: *Violence and warfare among hunter-gatherers*, ed. M. W. Allen & T. L. Jones, pp. 26–46. Routledge. [WB]

LeBlanc S.A. & Register K. (2003) *Constant battles: The myth of the peaceful, noble savage*. St. Martin's Press. [AM]

Lebreton M., Jorge S., Michel V., Thirion B. & Pessiglione M. (2009) An automatic valuation system in the human brain: Evidence from functional neuroimaging. *Neuron* **64**:431–39. [aCKWDD]

Ledgerwood A., Liviatan I. & Carnevale P. J. (2007) Group identity completion and the symbolic value of property. *Psychological Science* **18**:873–78. [rCKWDD]

Lerner J. S. & Keltner D. (2000) Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion* **14**(4):473–93. Available at: https://doi.org/10.1080/026999300402763. [PCC]

Lerner J. S. & Keltner D. (2001) Fear, anger, and risk. *Journal of Personality and Social Psychology* **81**(1):146–59. Available at: https://doi.org/10.1037/0022-3514.81.1.146. [PCC]

Levati M. V., Sutter M. & Van der Heijden R. (2007) Leading by example in a public goods experiment with heterogeneity and incomplete information. *Journal of Conflict Resolution* **51**:793–818. [aCKWDD]

Levy D. J. & Glimcher P. W. (2011) Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience* **31**:14693–707. [aCKWDD]

Levy D. J. & Glimcher P. W. (2012) The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology* **22**:1027–38. [aCKWDD]

Leyens J. P., Demoulin S., Vaes J., Gaunt R. & Paladino M. P. (2007) Infra-humanization: The wall of group differences. *Social Issues and Policy Review* **1**:139–72. [aCKWDD]

Li K., Szolnoski A., Cong R. & Wang L. (2016) The coevolution of overconfidence and bluffing in the resource competition game. *Scientific Reports* **6**:21104. [aCKWDD]

Li N. P., Lim A. J. Y., Tsai M.-H. & O J. (2015) Too materialistic to get married and have children? *PLoS One* **10**:e0126543. [JO]

Li X. & Camerer C. F. (2019) Using visual salience in empirical game theory. Working paper, *CalTech. [rCKWDD]*

Liberman P. & Skitka L. J. (2017) Revenge in US Public Support for War against Iraq. *Public Opinion Quarterly* **81**(3):636–60. [TMA]

Liu N., Hadj-Bouziane F., Jones K. B., Turchi J. N., Averbeck B. B. & Ungerleider L. G. (2015) Oxytocin modulates fMRI responses to facial expression in macaques. *Proceedings of the National Academy of Sciences USA* **112**:e3123–30. Available at: https://www.pnas.org/content/112/24/E3123.long. [RH]

Loerakker B. & Van Winden F. (2017) Emotional leadership in an intergroup conflict game. *Journal of Economic Psychology* **63**:143–67. [aCKWDD]

Lopes B. & Jaspal R. (2015) Paranoia predicts out-group prejudice: Preliminary experimental data. *Mental Health, Religion & Culture* **18**(5):380–95. [DK]

Lopez A. C. (2010) Evolution, coalitional psychology, and war (No. I). H-Diplo ISSF Roundtable on "Biology and Security." [ACL]

Lopez A. C. (2016) Conditions required for evolution of warfare adaptations. In: *Encyclopedia of evolutionary psychological science*, ed. V. Weekes-Shackelford, T. K. Shackelford & V. A. Weekes-Shackelford, pp. 1–10. Springer International. Available at: https://doi.org/10.1007/978-3-319-16999-6_914-1. [ACL]

Lopez A. C. (2017) The evolutionary psychology of war: Offense and defense in the adapted mind. *Evolutionary Psychology* **15**(4):1474704917742720. Available at: https://doi.org/10.1177/1474704917742720. [aCKWDD, ACL]

Lykken D. T. (1957) A study of anxiety in the sociopathic personality. *The Journal of Abnormal and Social Psychology* **55**(1):6–10. [TOP]

Macfarlan S. J., Erickson P. I., Yost J., Regalado J., Jaramillo L. & Beckerman S. (2018) Bands of brothers and in-laws: Waorani warfare, marriage and alliance formation. *Proceedings of the Royal Society B: Biological Sciences* **285**(1890). pii: 20181859. doi: 10.1098/rspb.2018.1859. [WB]

Macfarlan S. J., Walker R. S., Flinn M. V. & Chagnon N. A. (2014) Lethal coalitionary aggression and long-term alliance formation among Yanomamö men. *Proceedings of the National Academy of Sciences USA* **111**(47):16662–69. [WB, aCKWDD]

Mackie D. M., Devos T. & Smith E. R. (2000) Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology* **79**(4):602–16. [TMA]

Mago S. D., Samak A. C. & Sheremeta R. M. (2016) Facing your opponents: Social identification and information feedback in contests. *Journal of Conflict Resolution* **60**:459–81. [RMS]

Majolo B., de Bortoli Vizoli A. & Lehmann J. (2016) The effect of intergroup competition on intragroup affiliation in primates. *Animal Behaviour* **114**:13–19. [ARR]

Manchester W. (1980) *Goodbye, darkness: A memoir of the Pacific war*. Little, Brown & Company. [aCKWDD]

Mann M. (2018) Have wars and violence declined? *Theory and Society* **47**:37–60. [aCKWDD]

Manson J. H. & Wrangham R. W. (1991) Intergroup aggression in chimpanzees and humans. *Current Anthropology* **32**(4):369–90. [WB]

Margie N. G., Killen M., Sinno S. & McGlothlin H. (2005) Minority children's intergroup attitudes about peer relationships. *British Journal of Developmental Psychology* **23**(2):251–70. doi: 10.1348/026151005X26075. [NMcL]

Marie A. & Fitouchi L. (in preparation) The evolution of moral rigidity. [AM]

Marsh A. A. (2018) The caring continuum: Evolved hormonal and proximal mechanisms explain prosocial and antisocial extremes. *Annual Reviews in Psychology* **70**:347–71. Available at: https://www.annualreviews.org/doi/full/10.1146/annurev-psych-010418-103010?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_ dat=cr_pub%3Dpubmed. [RH]

Marsh N., Scheele D., Feinstein J. S., Gerhardt H., Strang S., Maier W. & Hurlemann R. (2017) Oxytocin-enforced norm compliance reduces xenophobic outgroup rejection. *Proceedings of the National Academy of Sciences USA* **114**(35):9314–19. Available at: https://www.pnas.org/content/114/35/9314.long. [RH]

Marsh N., Scheele D., Gerhardt H., Strang S., Enax L., Weber B., Maier W. & Hurlemann R. (2015) The neuropeptide oxytocin induces a social altruism bias. *Journal of Neuroscience* **35**(47):15696–701. Available at: http://www.jneurosci.org/content/35/47/15696.long. [RH]

Martin C. F., Bhui R., Bossaerts P., Matsuzawa T. & Camerer C. (2014) Chimpanzee choice rates in competitive games match equilibrium game theory predictions. *Scientific Reports* **4**:5182. [MWK]

Masuda N. (2012) In-group favoritism and intergroup cooperation under intergroup reciprocity based on group reputation. *Journal of Theoretical Biology* **21**:8–18. [aCKWDD]

May D., Radar N. E. & Goodrum S. (2010) A gendered assessment of the "threat of victimization." *Criminal Justice Review* **35**(2):159–182. [DVB]

McBride M. & Skaperdas S. (2014) Conflict, settlement, and the shadow of the future. *Journal of Economic Behavior & Organization* **105**:75–89. [aCKWDD]

McClung J.S., Triki Z., Clément F., Bangerter A., Bshary R. (2018) Endogenous oxytocin predicts helping and conversation as a function of group membership. *Proceedings of the Royal Society B: Biological Sciences* **285**(1882). pii: 20180939. Available at: https://royalsocietypublishing.org/doi/full/10.1098/rspb.2018.0939?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed&. [RH]

McCusker C. & Carnevale P. J. (1995) Framing in resource dilemmas: Loss aversion and the moderating effects of sanctions. *Organizational Behavior and Human Decision Processes* **61**:190–201. [aCKWDD]

McDermott R. (2015) Sex and death: Gender differences in aggression and motivations for violence. *International Organization* **69**(3):753–75. Available at: https://doi.org/10.1017/S0020818315000065. [ACL]

McDonald M. M., Navarrete C. D. & Van Vugt M. (2012) Evolution and the psychology of intergroup conflict: The male warrior hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**(1589):670–79. Available at: https://doi.org/10.1098/rstb.2011.0301. [ACL]

McDonough F. (1997) *The origins of the First and Second World Wars*. Cambridge University Press. [aCKWDD]

McGregor I., Hayes J. & Prentice M. (2015) Motivation for aggressive religious radicalization: Goal regulation theory and a personality × threat × affordance hypothesis. *Frontiers in Psychology* **6**:1325. [DK]

McKay R., Efferson C., Whitehouse H. & Fehr E. (2011) Wrath of God: Religious primes and punishment. *Proceedings of the Royal Society B: Biological Sciences* **278**(1713):1858–63. [aCKWDD]

McLoughlin N. & Over H. (2017) Young children are more likely to spontaneously attribute mental states to members of their own group. *Psychological Science* **28**(10):1503–509. doi: 10.1177/0956797617710724. [NMcL]

McLoughlin N. & Over H. (2018) The developmental origins of dehumanization. *Advances in Child Development and Behavior* **54**:153–78. doi: 10.1016/bs.acdb.2017.10.006. [NM, NMcL]

McNamara J. M. (2013) Towards a richer evolutionary game theory. *Journal of the Royal Society Interface* **10**:20130544. [ANR]

McNamara J. M. & Houston A. I. (2005) If animals know their own fighting ability, the evolutionarily stable level of fighting is reduced. *Journal of Theoretical Biology* **232**:1–6. [ANR]

McNaughton N. (1989) *Biology and emotion*. Cambridge University Press. [NMcN]

McNaughton N. (2018) What do you mean "anxiety"? Developing the first anxiety syndrome biomarker. *Journal of the Royal Society of New Zealand* **48**:177–90. doi: 10.1080/03036758.2017.1358184. [NMcN]

McNaughton N. & Corr P. J. (2004) A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews* **28**:285–305. doi:10.1016/j.neubiorev.2004.03.005. [NMcN]

McNaughton N. & Corr P. J. (2018) Survival circuits and risk assessment. *Current Opinion in Behavioral Sciences* **24**:14–20. doi:10.1016/j.cobeha.2018.01.018. [NMcN]

McNaughton N., DeYoung C. G. & Corr P. J. (2016) Approach/avoidance. In: *Neuroimaging personality, social cognition and character*, ed. J. R. Absher & J. Cloutier, pp. 25–49. Elsevier. [NMcN]

Meadowcroft J. & Morrow E. A. (2017) Violence, self-worth, solidarity and stigma: How a dissident, far-right group solves the collective action problem. *Political Studies* **65**(2):373–90. [DK]

Mehta P. H. & Beer J. (2010) Neural mechanisms of the testosterone-aggression relations: The role of the orbitofrontal cortex. *Journal of Cognitive Neuroscience* **22**:2357–68. [aCKWDD]

Meloy J. R. & Gothard S. (1995) Demographic and clinical comparison of obsessional followers and offenders with mental disorders. *American Journal of Psychiatry* **152**:258–63. [aCKWDD]

Meltzoff A. N. (2007) 'Like me': A foundation for social cognition. *Developmental Science* **10**(1):126–34. doi: 10.1111/j.1467-7687.2007.00574.x. [NMcL]

Messick D. M. & Cook K. S. (1983) *Equity theory: Psychological and sociological perspectives*. Praeger. [KU]

Messick D. M. & Thorngate W. B. (1967) Relative gain maximization in experimental games. *Journal of Experimental Social Psychology* **3**:85–101. [aCKWDD]

Meyer-Lindenberg A., Domes G., Kirsch P., Heinrichs M. (2011) Oxytocin and vasopressin in the human brain: Social neuropeptides for translational medicine. *Nature Reviews Neuroscience* **12**:524–38. Available at: https://www.nature.com/articles/nrn3044. [RH]

Mgbeoji I. (2006) The civilised self and the barbaric other: Imperial delusions of order and the challenges of human security. *Third World Quarterly* **27**:855–69. [aCKWDD]

Mifune N., Hizen Y., Kamijo Y. & Okano Y. (2016) Preemptive striking in individual and group conflict. *PLoS One* **11**(5):e0154859. [NM]

Mifune N., Simunovic D. & Yamagishi T. (2017) Intergroup biases in fear-induced aggression. *Frontiers in Psychology* **8**:49. [NM]

Mikolic J. M., Parker J. C. & Pruitt D. G. (1997) Escalation in response to persistent annoyance: Groups versus individuals and gender effects. *Journal of Personality and Social Psychology* **72**:151–63. [aCKWDD]

Mill J. S. (1848/2008) *Principles of political economy*. Oxford University Press. [aCKWDD]

Miller B. (2009) Between revisionist and the frontier state: Regional variations in state war-propensity. *Review of International Studies* **35**:85–119. [aCKWDD]

Miller D. T. & Holmes J. G. (1975) The role of situational restrictiveness on self-fulfilling prophecies: A theoretical and empirical extension of Kelley and Stahelski's triangle hypothesis. *Journal of Personality and Social Psychology* **31**:661–73. [rCKWDD]

Minkov M., Dutt P., Schachner M., Morales O., Sanchez C., Jandosova J., Khassenbekov Y. & Mudd B. (2017) A revision of Hofstede's individualism-collectivism dimension: A new national index from a 56-country study. *Cross Cultural & Strategic Management* **24**(3):386–404. Available at: https://doi.org/10.1108/CCSM-11-2016-0197. [AF]

Mirville M. O. (2018) The causes and consequences of intergroup interactions in mountain gorillas (*Gorilla beringei beringei*). PhD thesis, University of Western Australia. [ARR]

Mirville M. O., Ridley A. R., Samedi J. P. M., Vecellio V., Ndagijimana F., Stoinski T. S. & Grueter C. C. (2018) Factors influencing individual participation during intergroup interactions in mountain gorillas. *Animal Behaviour* **144**:75–86. [ARR]

Misch A., Over H. & Carpenter M. (2016) I won't tell: Young children show loyalty to their group by keeping group secrets. *Journal of Experimental Child Psychology* **142**:96–106. doi: 10.1016/j.jecp.2015.09.016. [NMcL]

Mobbs D. & LeDoux J. E. (2018) Editorial overview: Survival behaviors and circuits. *Current Opinion in Behavioral Sciences* **24**:168–71. doi:10.1016/j.cobeha.2018.10.004. [NMcN]

Molenberghs P. (2013) The neuroscience of in-group bias. *Neuroscience and Biobehavioral Reviews* **8**:1530–36. [aCKWDD]

Molenberghs P., Trautwein F. M., Bockler A., Singer T. & Kanske P. (2016) Neural correlates of metacognitive ability and feeling confident: A large-scale fMRI study. *Social Cognitive and Affective Neuroscience* **11**:1942–51. [aCKWDD]

Montoya E. R., Terburg D., Bos P. A. & van Honk J. (2012) Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and Emotion* **36**:65–73. [aCKWDD]

Motta S. C., Carobrez A. P. & Canteras N. S. (2017) The periaqueductal gray and primal emotional processing critical to influence complex defensive responses, fear learning and reward seeking. *Neuroscience and Biobehavioral Reviews* **76**:39–47. doi:10.1016/j.neubiorev.2016.10.012. [NMcN]

Mroszczyk J. (2016) To die or to kill? An analysis of suicide attack lethality. *Terrorism and Political Violence* **31**(2):346–366. [DK]

Mulder L. B., Van Dijk E., De Cremer D. & Wilke H. A. M. (2006) When sanctions fail to increase cooperation in social dilemmas: Considering the presence of an alternative option to defect. *Personality and Social Psychology Bulletin* **32**:1312–24. [rCKWDD]

Mummendey A., Kessler T., Klink A. & Mielke R. (1999) Strategies to cope with negative social identity: Predictions by social identity theory and relative deprivation theory. *Journal of Personality and Social Psychology* **76**:229–45. [NS]

Nadler A. & Shnabel N. (2015) Intergroup reconciliation: Instrumental and socio-emotional processes and the need based model. *European Review of Social Psychology* **26**:93–125. [NS]

Nakashima N. A., Halali E. & Halevy N. (2017) Third parties promote cooperative norms in repeated interactions. *Journal of Experimental Social Psychology* **68**:212–23. [aCKWDD]

Neale M. A. & Bazerman M. H. (1985) The effects of framing and overconfidence on bargaining behaviors and outcomes. *Academy of Management Journal* **28**:34–49. [aCKWDD]

Nelson L. D. & Foell J. (2018) Externalizing proneness and psychopathy. In: *Handbook of psychopathy*, ed. C. J. Patrick, 2nd ed., pp. 127–43. Guilford Press. [TOP]

Nelson R. J. & Trainor B. C. (2007) Neural mechanisms of aggression. *Nature Reviews Neuroscience* **8**:536–46. Available at: https://doi.org/10.1038/nrn2174. [PCC, aCKWDD]

Neuberg S. L. & Schaller M. (2016) An evolutionary threat-management approach to prejudices. *Current Opinion in Psychology* **7**:1–5. [DK]

Neuberg S. L., Warner C. M., Mistler S. A., Berlin A., Hill E. D., Johnson J. D., Filip-Crawford G., Millsap R. E., Thomas G., Winkelman M., Broome B. J., Taylor T. J. & Schober J. (2014) Religion and intergroup conflict: Findings from the Global Group Relations Project. *Psychological Science* **25**:198–206. [aCKWDD]

Nichols S. & Folds-Bennett T. (2003) Are children moral objectivists? Children's judgments about moral and response-dependent properties. *Cognition* **90**(2):B23–B32. Available at: http://dx.doi.org/10.1016/s0010-0277(03)00160-4. [AM]

Norenzayan A. & Shariff A. F. (2008) The origin and evolution of religious pro-sociality. *Science* **322**:58–62. [aCKWDD]

Norenzayan A., Shariff A. F., Gervais W. M., Willard A. K., McNamara R. A., Slingerland E. & Henrich J. (2016) The cultural evolution of prosocial religions. *Behavioral and Brain Sciences* **39**:1–65. [aCKWDD]

Nosenzo D., Offerman T., Sefton M. & van der Veen A. (2013) Encouraging compliance: Bonuses versus fines in inspection games. *Journal of Law, Economics, and Organization* **30**:623–48. [aCKWDD, MWK]

Novemsky N. & Kahneman D. (2005) The boundaries of loss aversion. *Journal of Marketing Research* **42**:119–28. [NMcN]

Nowak M. A. (2006) Five rules for the evolution of cooperation. *Science* **314**:1560–63. [aCKWDD]

Nowak M. A., Tarnita C. E. & Wilson E. O. (2010) The evolution of eusociality. *Nature* **466**:1057–62. [aCKWDD]

Nti K. O. (1999) Rent-seeking with asymmetric valuations. *Public Choice* **98**(3/4):415–30. [MWK]

Nunn C.L. (2000) Collective benefits, free-riders and male extra-group conflict. In: *Primate males*, ed P. M. Kappeler, pp. 192–204, Cambridge University Press. [ARR]

O J. (2018a) Learned helplessness from an evolutionary mismatch perspective. In: *Encyclopedia of evolutionary psychological science*, ed. T. K. Shackelford & V. A. Weekes-Shackelford. Springer Nature. [JO]

O J. (2018b) Self-efficacy, animal phobias and evolutionary mismatch. In: *Encyclopedia of evolutionary psychological science*, ed. T. K. Shackelford & V. A. Weekes-Shackelford. Springer Nature. [JO]

Oakley B. A. (2013) Concepts and implications of altruism bias and pathological altruism. *Proceedings of the National Academy of Sciences USA* **110**(Suppl. 2):10408–15. Available at: https://www.pnas.org/content/110/Supplement_2/10408.long. [RH]

Oka R. C., Kissel M., Golitko M., Sheridan S. G., Kim N. C. & Fuentes A. (2017) Population is the main driver of war group size and conflict casualties. *Proceedings of the National Academy of Sciences USA* **114**:11101–10. [aCKWDD]

Oprea R., Charness G. & Friedman D. (2014) Continuous time and communication in a public-goods experiment. *Journal of Economic Behavior and Organization* **108**:212–23. [aCKWDD]

Orbell J. M. & Dawes R. M. (1993) Social welfare, cooperators advantage, and the option of not playing the game. *American Sociological Review* **58**:787–800. [rCKWDD]

Osgood J. M. (2017) Is revenge about retributive justice, deterring harm, or both? *Social and Personality Psychology Compass* **11**(1):1–15. [DK]

Otsubo H. (2015) Nash equilibria in a two-person discrete all-pay auction with unfair tie-break and complete information. *Economics Bulletin* **35**:2443–54. [aCKWDD]

Otterbein K. (2009) *The anthropology of war*. Waveland Press. [WB]

Over H. (2016) The origins of belonging: Social motivation in infants and young children. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**:20150072. doi: 10.1098/rstb.2015.0072. [NMcL]

Over H. & McCall C. (2018) Becoming us and them: Social learning and intergroup bias. *Social and Personality Psychology Compass* **12**(4):e12384. doi: 10.1111/spc3.12384. [NMcL]

Oxley D. R., Smith K. B., Alford J. R., Hibbing M. V., Miller J. L., Scalora M., Hatemi P. K. & Hibbing J. R. (2008) Political attitudes vary with physiological traits. *Science* **321**:1667–70. [PP]

Palmer C. T. & Tilley C. F. (1995) Sexual access to females as a motivation for joining gangs: An evolutionary approach. *Journal of Sex Research* **32**:213–17. [DVB]

Patrick C. J. (1994) Emotion and psychopathy: Startling new insights. *Psychophysiology* **31**:319–30. [aCKWDD]

Patrick C. J. (2010) Operationalizing the triarchic conceptualisation of psychopathy: Preliminary description of brief scales for assessment of boldness, meanness, and disinhibition. *PhenX Toolkit Online Assessment Catalog*. Available at: https://www.phenxtoolkit.org/toolkit_content/supplemental_info/psychiatric/measures/Triarchic_Psychopathy_Measure_Manual.pdf. [TOP]

Patrick C. J. & Drislane L. E. (2015) Triarchic model of psychopathy: Origins, operationalizations, and observed linkages with personality and general psychopathology. *Journal of Personality* **83**(6):627–43. Available at: https://doi.org/10.1111/jopy.12119. [TOP]

Patrick C. J., Fowles D. C. & Krueger R. F. (2009) Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology* **21**(3):913–38. Available at: https://doi.org/10.1017/S0954579409000492. [TOP]

Paulus M. P. & Stein M. B. (2006) An insular view of anxiety. *Biological Psychiatry*, **60**(4):383–87. doi:10.1016/j.biopsych.2006.03.042. [NMcN]

Pehrson S., Devaney L., Blaylock D. & Bryan D. (2017) Beyond group engagement: Multiple pathways from encounters with the police to cooperation and compliance in Northern Ireland. *PLoS One* **12**(9):e0184436. Available at: https://doi.org/https://doi.org/10.1371/journal.pone.0184436. [KU]

Peterson C. K., Gable P. & Harmon-Jones E. (2008) Asymmetrical frontal ERPs, emotion, and behavioral approach/inhibition sensitivity. *Social Neuroscience* **3**:113–24. [aCKWDD]

Piazza J. & Sousa P. (2013) Religiosity, political orientation, and consequentialist moral yhinking. *Social Psychological and Personality Science* **5**(3):334–42. [AM]

Pietrazewski D. (2016) How the mind sees group and coalitionary conflict: The evolutionary invariances of n-person conflict dynamics. *Evolution and Human Behavior* **37**:470–80. [aCKWDD]

Pinker S. (2011) *The better angels of our mind*. Allen Lane. [aCKWDD]

Pinkley R. L. (1995) Impact of knowledge regarding alternatives to settlement in dyadic negotiations: Whose knowledge counts? *Journal of Applied Psychology* **80**:403–17. [rCKWDD]

Pliskin R. & Halperin E. (2016) Emotions and emotion regulation in intractable conflict and their relation to the ethos of conflict in Israeli society. In: *A social psychology perspective on the Israeli-Palestinian conflict*, ed. K. Sharvit & E. Halperin. Peace Psychology Book Series. Springer. [rCKWDD]

Plous S. (1985) Perceptual illusions and military realities: The nuclear arms race. *Journal of Conflict Resolution* **29**:363–89. [arCKWDD, NH]

Porat R., Halperin E. & Tamir M. (2016) What we want is what we get: Group-based emotional preferences and conflict resolution. *Journal of Personality and Social Psychology* **110**(2):167–90. Available at: https://doi.org/10.1037/pspa0000043. [PCC]

Posner M. I. & Rothbart M. K. (2007) Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology* **58**:1–23. [aCKWDD]

Potegal M. (2012) Temporal and frontal lobe initiation and regulation of the top-down escalation of anger and aggression. *Behavioral Brain Research* **231**:386–95. [aCKWDD]

Potters J., Sefton M. & Vesterlund L. (2007) Leading-by-example and signaling in voluntary contribution games: An experimental study. *Economic Theory* **33**:169–82. [aCKWDD]

Pratto F. & John O. P. (1991) Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology* **61**:380–91. [aCKWDD]

Pratto F., Sidanius J. & Levin S. (2006) Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology* **17**:271–320. [DVB]

Prediger S., Vollan B. & Benedikt H. (2014) Resource scarcity and antisocial behavior. *Journal of Public Economics* **119**:1–9. [aCKWDD]

Pruitt D. G. (1967) Reward structure and cooperation: The decomposed Prisoner's Dilemma Game. *Journal of Personality and Social Psychology* **7**:21–27. [aCKWDD]

Pruitt D. G. (1981) *Negotiation*. Academic Press. [arCKWDD]

Pruitt D. G. (1998) Social conflict. In: *Handbook of social psychology*, ed. D. Gilbert, S. T. Fiske & G. Lindzey, 4th ed., vol. **2**, pp. 89–150. McGraw-Hill. [aCKWDD]

Pruitt D.G. (2007) Readiness theory and the Northern Ireland conflict. *American Behavioral Scientist* **50**:1520–41. [rCKWDD]

Pruitt D. G. & Kimmel M. J. (1977) Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology* **28**(1):363–92. [aCKWDD, NM]

Pruitt D. G. & Rubin J. Z. (1986) *Social conflict: Escalation, stalemate, and settlement*. Random House. [arCKWDD]

Purzycki B. B., Apicella C., Atkinson Q. D., Cohen E., McNamara R. A., Willard A. K., Xygalatas D., Norenzayan A. & Henrich J. (2016) Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature* **530**:327. [aCKWDD]

Putnam R. D. (1988) Diplomacy and domestic politics: The logic of two-level games. *International Organization* **42**(3):427–60. [NH]

Quillian L. (1995) Prejudice as a response to perceived group threat – Population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review* **60**:586–611. [aCKWDD]

Quine W. V. (1948) On what there is. *The Review of Metaphysics* **2**(1):21–38. [AM]

Radburn M. & Stott C. (2018) The social psychological processes of "procedural justice": Concepts, critiques and opportunities. *Criminology and Criminal Justice*. Available at: https://doi.org/10.1177/1748895818780200. [KU]

Radburn M., Stott C., Bradford B. & Robinson M. (2016) When is policing fair? Groups, identity and judgements of the procedural justice of coercive crowd policing. *Policing and Society* **28**(6):647–664. Available at: https://doi.org/10.1080/10439463.2016.1234470. [KU]

Radcliffe-Brown A. R. (1922) *The Andaman islanders: A study in social anthropology*. Cambridge University Press. [WB]

Radford A. N. (2003) Territorial vocal rallying in the green woodhoopoe: influence of rival group size and composition. *Animal Behaviour* **66**:1035–44. [ANR]

Radford A. N. (2005) Neighbour-stranger discrimination in the group-living green woodhoopoe. *Animal Behaviour* **70**:1227–34. [ANR]

Radford A. N. (2008) Duration and outcome of intergroup conflict influences intragroup affiliative behavior. *Proceedings of the Royal Society B: Biological Sciences* **275**:2787–91. [aCKWDD, ARR]

Radford A. N. (2011) Preparing for battle? Potential intergroup conflict promotes current intragroup affiliation. *Biology Letters* **7**:26–29. [ANR]

Radford A. N. & du Plessis M. A. (2004) Territorial vocal rallying in the green woodhoopoe: Factors affecting the contest length and outcome. *Animal Behaviour* **68**:803–10. [ANR]

Radford A. N., Majolo B. & Aureli F. (2016) Within-group behavioural consequences of between-group conflict: A prospective review. *Proceedings of the Royal Society B: Biological Sciences* **283**(1843):20161567. [aCKWDD, ANR]

Rai T. S. & Fiske A. P. (2011) Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review* **118**:57–75. [aCKWDD]

Rai T. S., Valdesolo P. & Graham J. (2017) Dehumanization increases instrumental violence, but not moral violence. *Proceedings of the National Academy of Sciences USA* **114**:8511–16. [aCKWDD]

Raiffa H. (1982) *The art and science of negotiation*. Belknap. [rCKWDD]

Raihani N. J. & Bell V. (2018) An evolutionary perspective on paranoia. *Nature Human Behaviour* **3**:114–121. [DK]

Raine A. & Yang Y. (2006) Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience* **1**(3):203–13. Available at: https://doi.org/10.1093/scan/nsl033. [PCC]

Raleigh C. & Hegre H. (2009) Population size, concentration, and civil war: A geographically disaggregated analysis. *Political Geography* **28**:224. [aCKWDD]

Rand D. G., Greene J. D. & Nowak M. A. (2012) Spontaneous giving and calculated greed. *Nature* **489**:427–30. [aCKWDD]

Rapoport A. (1960) *Fights, games, and debates*. Michigan University Press. [aCKWDD]

Rapoport A. & Bornstein G. (1987) Intergroup competition for the provision of binary public-goods. *Psychological Review* **94**:291–99. [aCKWDD]

Rauhut H. (2009) Higher punishment, less control? Experimental evidence on the inspection game. *Rationality and Society* **21**(3):359–92. [MWK]

Rauthmann J. F., Gallardo-Pujol D., Guillaume E. M., Todd E., Nave C. S., Sherman R. A., Ziegler M., Jones A. B. & Funder D. C. (2014) The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology* **107**:677. [rCKWDD]

Reicher S. D. (1996) "The Battle of Westminster": Developing the social identity model of crowd behaviour in order to explain the initiation and development of collective conflict. *European Journal of Social Psychology* **26**:115–34. Available at: https://doi.org/10.1002/(sici)1099-0992(199601)26:1<115::aid-ejsp740>3.3.co;2-q. [KU]

Reicher S. D., Haslam S. A. & Smith J. R. (2012) Working toward the experimenter: Reconceptualizing obedience within the Milgram paradigm as identification-based followership. *Perspectives on Psychological Science* **7**:315–24. [PP]

Renfrew J. W. & Hutchinson R. R. (1983) The motivation of aggression. In E. Satinoff & P. Teitelbaum (Eds.), *Motivation* (Vol. **6**): Plenum Press. [NMcN]

Renno M. P. & Shutts K. (2015) Children's social category-based giving and its correlates: expectations and preferences. *Developmental Psychology* **51**(4):533–43. doi: 10.1037/a0038819. [NMcL]

Rhodes M. & Mandalaywala T. M. (2017) The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science* **8**(4):1437. doi:10.1002/wcs.1437. [NMcL]

Ridley A. R. (2016) Southern pied babblers: the dynamics of conflict and cooperation in a group-living society. In: *Cooperative breeding in vertebrates: studies in ecology, evolution and behaviour*, ed. W. D. Koenig & J. L. Dickinson, pp. 115–32. Cambridge University Press. [ARR]

Rilling J. K. & Sanfey A. G. (2011) The neuroscience of social decision making. *Annual Review of Psychology* **62**:23–48. [aCKWDD]

Robert C. & Carnevale P. J. (1997) Group choice in ultimatum bargaining. *Organizational Behavior and Human Decision Processes* **72**:256–79. doi:10.1006/obhd.1997.2738. [JH]

Robinson R. J. & Keltner D. (1996) Much ado about nothing? Revisionists and traditionalists choose an introductory English syllabus. *Psychological Science* **7**:18–24. [aCKWDD]

Roccas S., Sagiv L. & Schwartz S. (2008) Toward a unifying model of identification with groups. Integrating theoretical perspectives. *Personality and Social Psychology Review* **12**:280–306. [aCKWDD]

Ronay R. & von Hippel W. (2010) The presence of an attractive women elevates testosterone and risk taking in young men. *Social Psychological and Personality Science* **1**:57–64. [DVB]

Roskes M., Elliot A. & De Dreu C. K. W. (2014) Regulating avoidance motivation: A conservation of energy approach. *Current Directions in Psychological Science* **23**:133–38. [aCKWDD]

Ross L. & Ward A. (1995) Psychological barriers to dispute resolution. *Advances in Experimental Social Psychology* **27**:255–304. San Diego: Academic Press. [aCKWDD]

Rucker D. D., Galinsky A. D. & Magee J. C. (2018) The agentic-communal model of advantage and disadvantage: How inequality produces similarities in the psychology

of power, social class, gender, and race. In: *Advances in Experimental Social Psychology (vol. 58)*, ed. J. M. Olson & M. P. Zanna, pp. 71–125. Academic Press. [NS]

Rusch H. (2013) Asymmetries in altruistic behavior during violent intergroup conflict. *Evolutionary Psychology* **11**(5):973–93. [TMA, WB, aCKWDD, ACL]

Rusch H. (2014a) The two sides of warfare: An extended model of altruistic behavior in ancestral human intergroup conflict. *Human Nature* **25**(3):359–77. https://doi.org/10.1007/s12110-014-9199-y [WB, aCKWDD, ACL, HR]

Rusch H. (2014b) The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension. *Proceedings of the Royal Society B: Biological Sciences* **281**(1794):20141539. [TMA, aCKWDD]

Rusch H. & Gavrilets S. (2019) The logic of animal intergroup conflict: A review. *Journal of Economic Behavior & Organization*. Available at: https://doi.org/10.1016/j.jebo.2017.05.004. [aCKWDD, RMS]

Saalveld V., Ramadan Z., Bell V. & Raihani N. J. (2018) Experimentally induced social threat increases paranoid thinking. *Royal Society Open* **5**:180569. [aCKWDD]

Saguy T., Dovidio J. F. & Pratto F. (2008) Beyond contact: Intergroup contact in the context of power relations. *Personality and Social Psychology Bulletin* **34**:432–45. [NS]

Salter F. (2008) Evolutionary analyses of ethnic solidarity: An overview. *People and Place* **16**(2):41–51. [DS]

Samuelson W. & Zeckhauser R. (1988) Status quo bias in decision making. *Journal of Risk and Uncertainty* **1**:7–59. [ aCKWDD]

Sand H., Wikenros C., Wabakken P. & Liberg O. (2006) Effects of hunting group size, snow depth and age on the success of wolves hunting moose. *Animal Behavior* **72**:781–89. [aCKWDD]

Sander D., Grandjean D. & Scherer K. R. (2018) An appraisal-driven componential approach to the emotional brain. *Emotion Review* **10**(3):219–31. Available at: https://doi.org/10.1177/1754073918765653. [PCC]

Sapolsky R. M. (2005) The influence of social hierarchy on primate health. *Science* **308**:648–52. [aCKWDD]

Sapolsky R. M. (2017) *Behave*. Penguin. [a*CKWDD*]

Sapolsky R. M., Romero L. M. & Munck A. U. (2000) How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews* **21**:55–89. [aCKWDD]

Schaub M. (2017) Threat and parochialism in intergroup relations: Lab-in-the-field evidence from rural Georgia. *Proceedings of the Royal Society B: Biological Sciences* **284** (1865):20171560. [aCKWDD]

Scheele D., Striepens N., Kendrick K. M., Schwering C., Noelle J., Wille A., Schläpfer T. E., Maier W. & Hurlemann R. (2014) Opposing effects of oxytocin on moral judgment in males and females. *Human Brain Mapping* **35**(12):6067–76. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/hbm.22605. [RH]

Schelling T. C. (1960) *The strategy of conflict*. Harvard University Press. [aCKWDD]

Schelling T. C. (1980) *The strategy of conflict*. Harvard University Press. [NH]

Scherer K. R. & Moors A. (2019) The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology* **70**(1):719–45. Available at: https://doi.org/10.1146/annurev-psych-122216-011854. [PCC]

Schindler S. & Radford A. N. (2018) Factors influencing within-group conflict over defence against conspecific outsiders seeking breeding positions. *Proceedings of the Royal Society B: Biological Sciences* **285**:20181669. [ANR, ARR]

Schopler J., Insko C. A., Graetz K. A., Drigotas S., Smith V. A. & Dahl K. (1993) Individual-group discontinuity: Further evidence for mediation by fear and greed. *Personality and Social Psychology Bulletin* **19**:419–31. doi:10.1177/0146167293194007. [JH]

Segall G., Birnbaum D., Deeb I. & Diesendruck G. (2015) The intergenerational transmission of ethnic essentialism: How parents talk counts the most. *Developmental Science* **18**(4):543–55. doi: 10.1111/desc.12235. [NMcL]

Shadli S. M., McIntosh J., Glue P. & McNaughton N. (2015) An improved human anxiety process biomarker: Characterisation of frequency band, personality, and pharmacology. *Translational Psychiatry* **5**:e699. doi:10.1038/tp.2015.188. [NMcN]

Sharkin B. S. (2004) Road rage: Risk factors, assessment, and intervention strategies. *Journal of Counseling & Development* **82**(2):191–98. [DK]

Shaver P. R., Segev M. & Mikulincer M. (2011) A behavioral systems perspective on power and aggression. In: *Herzilya series on personality and social psychology. Human aggression and violence: Causes, manifestations, and consequences*, ed. P. R. Shaver & M. Mikulincer, pp. 71–87. American Psychological Association. [DK]

Shepherd L., Fasoli F., Pereira A. & Branscombe N. (2018) The role of threat, collective angst, and prejudice in promoting collective action against immigrant groups. *European Journal of Social Psychology* **48**(4):447–59. [PP]

Sheremeta R. M. (2010) Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior* **68**:731–47. [RMS]

Sheremeta R. M. (2018) Behavior in group contests: A review of experimental research. *Journal of Economic Surveys* **32**:683–704. [RMS]

Shutts K., Kinzler K. D., Katz R. C., Tredoux C. & Spelke E. S. (2011) Race preferences in children: Insights from South Africa. *Developmental Science* **14**(6):1283–91. doi: 10.1111/j.1467-7687.2011.01072.x. [NMcL]

Sidanius J. & Pratto F. (1999) *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press. [NS]

Siegel A., Roeling T. A. P., Gregg T. R. & Kruk M. R. (1999) Neuropharmacology of brain-stimulation-evoked aggression. *Neuroscience and Biobehavioral Reviews* **23**:359–98. [aCKWDD]

Silk J. B. & House B. R. (2011) Evolutionary foundations of human prosocial sentiments. *Proceedings of the National Academy of Sciences USA* **108**(Suppl. 2):10910–17. Available at: https://www.pnas.org/content/108/Supplement_2/10910.long. [RH]

Simandan D. (2010) Beware of contingency. *Environment and Planning. D: Society and Space* **28**(3):388–96. Available at: https://doi.org/10.1068/d2310. [DS]

Simandan D. (2018a) Competition, contingency, and destabilization in urban assemblages and actor-networks. *Urban Geography* **39**(5):655–66; https://doi.org/10.1080/02723638.2017.1382307. [DS]

Simandan D. (2018b) Wisdom and foresight in Chinese thought: Sensing the immediate future. *Journal of Futures Studies* **22**(3):35–50. Available at: https://doi.org/10.6531/JFS.2018.22(3).00A35. [DS]

Simandan D. (2019a) Being surprised and surprising ourselves: a geography of personal and social change. *Progress in Human Geography*. Available at: https://doi.org/10.1177/0309132518810431. [DS]

Simandan D. (2019b) Competition, delays, and coevolution in markets and politics. *Geoforum* **98**:15–24. Available at: https://doi.org/10.1016/j.geoforum.2018.09.014. [DS]

Simandan D. (2019c) Iterative lagged asymmetric responses in strategic management and long-range planning. *Time & Society*. Available at: https://doi.org/10.1177/0961463X17752652. [DS]

Simandan D. (2019d) Revisiting positionality and the thesis of situated knowledge. *Dialogues in Human Geography* **9**(2):129–49. Available at: https://doi.org/10.1177/2043820619850013. [DS]

Simon B. & Brown R. (1987) Perceived intragroup homogeneity in minority–majority contexts. *Journal of Personality and Social Psychology* **53**:703–11. [NS]

Simon H. A. (1956) Rational choice and the structure of the environment. *Psychological Review* **63**:129–38. [aCKWDD]

Simunovic D., Mifune N. & Yamagishi T. (2013) Preemptive strike: An experimental study of fear-based aggression. *Journal of Experimental Social Psychology* **49** (6):1120–3. [aCKWDD, NM]

Skaperdas S. & Syropoulos C. (1996) Can the shadow of the future harm cooperation? *Journal of Economic Behavior & Organization* **29**:355–72. [aCKWDD]

Skinner A. L. & Meltzoff A. N. (2019) Childhood Experiences and Intergroup Biases among Children. *Social Issues and Policy Review* **13**(1):211–40. doi: 10.1111/sipr.12054. [NMcL]

Slantchev B. (2010) Feigning weakness. *International Organization* **64**:357–88. [aCKWDD]

Small D. A., Gelfand M., Babcock L. & Gettman H. (2007) Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation. *Journal of Personality and Social Psychology* **93**(4):600–13. [NH]

Smyth K., Feeney A., Eidson R. C. & Coley J. D. (2017) Development of essentialist thinking about religion categories in Northern Ireland (and the United States). *Developmental Psychology* **53**(3):475–96. doi: 10.1037/dev0000253. [NMcL]

Snyder G. H. & Diesing P. (1977) *Conflict among nations: Bargaining, decision making, and system structure in international crises*. Princeton University Press. [aCKWDD]

Somma A., Borroni S., Drislane L. E., Patrick C. J. & Fossati A. (2019) Modeling the structure of the triarchic psychopathy measure: Conceptual, empirical, and analytic considerations. *Journal of Personality Disorders*. Available at: https://doi.org/10.1521/pedi_2018_32_354. [TOP]

Sonnemans J., Schram A. & Offerman T. (1998) Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior and Organization* **34**:143–61. [aCKWDD]

Spence A. M. (1973) Job market signaling. *Quarterly Journal of Economics* **87**:355–74. [CH]

Spengler F. B, Scheele D., Marsh N., Kofferath C., Flach A., Schwarz S., Stoffel-Wagner B., Maier W. & Hurlemann R. (2017) Oxytocin facilitates reciprocity in social communication. *Social Cognitive and Affective Neuroscience* **12**(8):1325–33. Available at: https://academic.oup.com/scan/article/12/8/1325/3746633. [RH]

Stallen M., De Dreu C.K., Shalvi S., Smidts A. & Sanfey A.G. (2012) The herding hormone: oxytocin stimulates in-group conformity. *Psychological Science* **23**(11):1288–92. Available at: https://journals.sagepub.com/doi/full/10.1177/0956797612446026?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed. [RH]

Stallen M., Rossi F., Heijne A., Smidts A., De Dreu C.K., Sanfey A.G. (2018) Neurobiological mechanisms of responding to injustice. *Journal of Neuroscience* **38** (12):2944–54. Available at: http://www.jneurosci.org/content/38/12/2944.long. [RH]

Stanford K. (2018) The difference between ice cream and Nazis: Moral externalization and the evolution of human cooperation. *Behavioral and Brain Sciences* **41**:e95. [AM]

Staub E. (1996) Cultural societal roots of violence – The examples of genocidal violence and of contemporary youth violence in the United States. *American Psychologist* **51**:117–32. [aCKWDD]

Stavenhagen R., (2016) *Ethnic conflicts and the nation-state*. Springer. [DS]

Steele M. A., Halkin S. L., Smallwood P. D., McKenna T. J., Mitsopoulos K. & Beam M. (2008) Cache protection strategies of a scatter-hoarding rodent: Do tree squirrels engage in behavioral deception? *Animal Behavior* **75**:705–14. [aCKWDD]

Steffens N. & Haslam S. A. (2013) Power through 'us': Leaders' use of we-referencing language predicts election victory. *PLoS One* **8**:e77952. [PP]

Steinel W., De Dreu C. K. W., Ouwehand E. & Ramírez-Marín J. Y. (2009) When constituencies speak in multiple tongues: The relative persuasiveness of hawkish minorities in representative negotiation. *Organizational Behavior and Human Decision Processes* **109**(1):67–78. [NH]

Stenner K. (2005) *The authoritarian dynamic*. Cambridge University Press. [PP]

Stern C., West T. V., Jost J. T. & Rule N. O. (2014) "Ditto heads": Do conservatives perceive greater consensus within their ranks than liberals? *Personality and Social Psychology Bulletin* **40**:1162–77. [PP]

Sternberg R. J. (2003) A duplex theory of hate: Development and application to terrorism, massacres, and genocide. *Review of General Psychology* **7**:299–328. [aCKWDD]

Stockley P. & Campbell A. (2013) Female competition and aggression: Interdisciplinary perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**:20130073. [DVB]

Stott C. & Reicher S. (1998) How conflict escalates: The inter-group dynamics of collective football crowd "violence." *Sociology* **32**:353–77. [aCKWDD]

Strang S., Gerhardt H., Marsh N., Oroz Artigas S., Hu Y., Hurlemann R., Park S. Q. (2017) A matter of distance – The effect of oxytocin on social discounting is empathy-dependent. *Psychoneuroendocrinology* **78**:229–32. Available at: https://www.sciencedirect.com/science/article/pii/S030645301630868X?via%3Dihub. [RH]

Strang S., Gross J., Schuhmann T., Riedl A., Weber B. & Sack A. T. (2015) Be nice if you have to – The neurobiological roots of strategic fairness. *Social Cognitive and Affective Neuroscience* **10**(6):790–96. [aCKWDD]

Striepens N., Kendrick K.M., Maier W. & Hurlemann R. (2011) Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Frontiers in Neuroendocrinology* **32**:426–50. Available at: https://www.sciencedirect.com/science/article/pii/S0091302211000665?via%3Dihub. [RH]

Striepens N., Scheele D., Kendrick K.M., Becker B., Schäfer L., Schwalba K., Reul J., Maier W. & Hurlemann R. (2012) Oxytocin facilitates protective responses to aversive social stimuli in males. *Proceedings of the National Academy of Sciences USA* **109**(44):18144–49. Available at: https://www.pnas.org/content/109/44/18144.long. [RH]

Strong M. J., Sherman B. L. & Riehl C. (2018) Home field advantage, not group size, predicts outcomes of intergroup conflicts in a social bird. *Animal Behaviour* **143**:205–13. [ANR]

Sugiura H., Mifune N., Tsuboi S. & Yokota K. (2017) Gender differences in intergoup conflict: The effect of outgroup threat priming on social dominance orientation. *Personality and Individual Differences* **104**:262–65. [DVB]

Swann Jr. W. B., Jetten J., Gomez A., Whitehouse H. & Bastian B. (2012) When group membership gets personal: A theory of identity fusion. *Psychological Review* **119** (3):441–56. [DK]

Tamir M. (2009) What do people want to feel and why? Pleasure and utility in emotion regulation. *Current Directions in Psychological Science* **18**(2):101–105. https://doi.org/10.1111/j.1467-8721.2009.01617.x [PCC]

Tamir M. (2016) Why do people regulate their emotions? A taxonomy of motives in emotion regulation. *Personality and Social Psychology Review* **20**(3):199–222. Available at: https://doi.org/10.1177/1088868315586325. [PCC]

Tankard M. E. & Paluck E. L. (2016) Norm perception as a vehicle for social change. *Social Issues and Policy Review* **10**:181–211. [KU]

Taylor S. E. (1991) Asymmetrical effects of positive and negative events – The mobilization minimization hypothesis. *Psychological Bulletin* **110**:67–85. [aCKWDD]

Ten Velden F. S., Beersma B. & De Dreu C. K. W. (2011) When competition breeds equality: Effects of appetitive versus aversive competition in negotiation. *Journal of Experimental Social Psychology* **47**:1127–33. [aCKWDD]

Ten Velden F. S., Daughters K. & De Dreu C. K. W. (2017) Oxytocin promotes intuitive rather than deliberated cooperation with the in-group. *Hormones and Behavior* **92**:164–71. [rCKWDD]

Tetlock P. E. (2003) Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences* **7**:320–24. [AM]

Tetlock P. E., Kristel O. V., Elson B., Green M. & Lerner J. (2000) The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology* **78**:853–70. [AM]

Thompson F. J., Marshall H. H., Vitikainen E. I. K. & Cant M. A. (2017) Causes and consequencs of intergroup conflict in cooperative banded mongooses. *Animal Behaviour* **126**:31–40. [ANR, ARR]

Tom S. M., Fox C. R., Trepel C. & Poldrack R. A. (2007) The neural basis of loss aversion in decision-making under risk. *Science* **315**(5811):515–18. https://doi.org/10.1126/science.1134239. [TOP]

Tooby J. & Cosmides L. (1988) The evolution of war and its cognitive foundations. Institute for Evolutionary Studies Technical Report 88–1. Available at: https://pdfs.semanticscholar.org/7f95/d9d117721df9e69b929b004d9d85ea6c560d.pdf?_ga=2.101426625. [ACL]

Tooby J. & Cosmides L. (1990) The past explains the present – Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology* **11**:375–424. [aCKWDD]

Tooby J. & Cosmides L. (2010) Groups in mind: The coalitional roots of war and morality. In: *Human morality and sociality: Evolutionary and comparative perspectives*, ed. H. Høgh-Olesen, pp. 191–234. Palgrave MacMillan. [AM]

Topalli V. & O'Neal E. C. (2003) Retaliatory motivation enhances attributions of hostility when people process ambiguous social stimuli. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression* **29**(2):155–72. [DK]

Traulsen A. & Nowak M. A. (2006) Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences USA* **103**:10952–55. [aCKWDD]

Treptow K.W., ed. (1997) *A history of Romania*. Center for Romanian Studies. [DS]

Trivers R. L. (1971) The evolution of reciprocal altruism. *Quarterly Review of Biology* **46**:35–57. [AM]

Trivers R. L. (1972) Parental investment and sexual selection. In: *Sexual selection and the descent of man: 1871–1971*, ed. B. Campbell, pp. 136–79. Chicago: Aldine. [DVB]

Tullock G. (1980) Efficient rent seeking. In: *Toward a theory of the rent-seeking society*, ed. J. M. Buchanan, R. D. Tollison & G. Tullock, pp. 97–112. Texas A&M University Press. [SMC, aCKWDD]

Turchin P. (2007) *War and peace and war: The rise and fall of empires*. Penguin. [AF]

Turner J. C. (2005) Explaining the nature of power: A three process theory. *European Journal of Social Psychology* **35**(1):1–22. [KU]

Tversky A. & Kahneman D. (1991) Loss aversion in riskless choice: A reference dependent model. *The Quarterly Journal of Economics* **106**:1039–61. [NMcN]

Tyler T. R. (2006) *Why people obey the law*, 2nd edition. Yale University Press. [KU]

Tyler T. R. & Lind E. A. (1992) A relational model of authority in groups. In: *Advances in experimental social psychology*, vol. **25**, pp. 115–82. Academic Press. Available at: https://doi.org/10.1016/S0065-2601(08)60283-X. [KU]

Ufkes E. G., Giebels E., Otten S. & Van der Zee K. I. (2014) The effectiveness of a mediation program in symmetrical versus asymmetrical neighbor-to-neighbor conflicts. *International Journal of Conflict Management* **23**:440–57. [aCKWDD]

Ule A., Schram A., Riedl A. & Cason T. N. (2009) Indirect punishment and generosity toward strangers. *Science* **326**:1701–704. [aCKWDD]

Valliancourt T. (2013) Do human females use indirect aggression as an intrasexual competitive strategy? *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**:20130080. [DVB]

Van de Vliert E. (1992) Questions about the strategic choice model of mediation. *Negotiation Journal* **8**:379–86. [aCKWDD]

Van de Vliert E. (2013) Climato-economic habitats support patterns of human needs, stresses, and freedoms. *Behavioral and Brain Sciences* **36**(5):465–480. [aCKWDD]

Van Dijk E., De Kwaadsteniet E. W. & De Cremer D. (2009) Tacit coordination in social dilemmas: The importance of having a common understanding. *Journal of Personality and Social Psychology* **96**:665–78. [aCKWDD]

Van Dijk E., Wilke H. & Wit A. (2003) Preferences for leadership in social dilemmas: Public good dilemmas versus common resource dilemmas. *Journal of Experimental Social Psychology* **39**:170–76. [aCKWDD]

Van Evera S. (2003) Why states believe foolish ideas: Non-self evaluation by states and societies. In: *Perspectives on structural realism*, ed. A. K. Hanami, pp. 163–98. Palgrave Macmillan. [aCKWDD]

Van Kleef G. A., Steinel W. & Homan A. C. (2013) On being peripheral and paying attention: Prototypicality and information processing in intergroup conflict. *Journal of Applied Psychology* **98**(1):63. [NH]

Van Lange P. A. M. (1999) The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientations. *Journal of Personality and Social Psychology* **77**:337–49. [aCKWDD]

Van Vugt M. (2009) Sex differences in intergroup competition, aggression, and warfare: The male warrior hypothesis. *Annals of the New York Academy of Sciences* **167**:124–34. Available at: https://doi.org/10.1111/j.1749-6632.2009.04539.x. [DVB, ACL]

Van Vugt M. & De Cremer D. (1999) Leadership in social dilemmas: The effects of group identification on collective actions to provide public goods. *Journal of Personality and Social Psychology* **76**:587–99. [aCKWDD]

Van Vugt M., De Cremer D. & Janssen D. P. (2007): Gender differences in cooperation and competition: The male warrior hypothesis. *Psychological Science* **18**:19–23. [DVB]

Van Zomeren M., Postmes T. & Spears R. (2008) Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin* **134**(4):504. [TMA]

Vermeij G. J. (1982) Unsuccessful predation and evolution. *American Naturalist* **120**:701–20. [aCKWDD]

Vogel D. L. & Karney B. R. (2002) Demands and withdrawal in newlyweds: Elaborating on the social structure hypothesis. *Journal of Social and Personal Relationships* **19**:685–701. [aCKWDD]

Von Clausewitz G. (1832/1984) *Von Kriege* (transl. M. Howard). Princeton University Press. [rCKWDD]

Von Neumann J. (1953) A certain zero-sum two-person game equivalent to the optimal assignment problem. In: *Contributions to the theory of games, vol. II*, ed. H. W. Kuhn & A. W. Tucker, pp. 5–12. Princeton University Press. [aCKWDD]

Wainryb C., Shaw L., Langley M., Cottam K. & Lewis R. (2004) *Children*'s thinking about diversity of belief in the early school years: Judgments of relativism, tolerance, and disagreeing persons. *Child Development* **75**(3):687–703. [AM]

Walker R. H., King A. J., McNutt J. W. & Jordan N. R. (2017) Sneeze to leave: African wild dogs (*Lycaon pictus*) use variable quorum thresholds facilitated by sneezes in collective decisions. *Proceedings of the Royal Society B: Biological Sciences* **284**:20170347. [aCKWDD]

Walker R. S. & Bailey D. H. (2013) Body counts in lowland South American violence. *Evolution and Human Behavior* **34**(1):29–34. Available at: https://doi.org/10.1016/j.evolhumbehav.2012.08.003. [WB, HR]

Walton R. E. & McKersie R. (1965) *A behavioral theory of labor negotiations: An analysis of a social interaction system*. McGraw-Hill. [rCKWDD]

Watson-Jones R. E. & Legare C. H. (2016) The social functions of rituals. *Current Directions in Psychological Science* **25**:42–46. [aCKWDD]

Waytz A., Young L. L. & Ginges J. (2014) Motive attribution asymmetry for love vs. hate drives intractable conflict. *Proceedings of the National Academy of Sciences USA* **111** (44):15687–92. [aCKWDD, DK]

Webster D. (1975) Warfare and the evolution of the state: A reconsideration. *American Antiquity* **40**:464–70. [aCKWDD]

Weinstein J. M. (2005) Resources and the information problem in rebel recruitment. *Journal of Conflict Resolution* **49**:598–624. [aCKWDD]

Weisel O. & Zultan R. (2016) Social motives in intergroup conflict: Group identity and perceived target of threat. *European Economic Review* **90**:122–33. Available at: .https://doi.org/https://doi.org/10.1016/j.euroecorev.2016.01.004. [aCKWDD, OW]

West S. A., Griffin A. S. & Gardner A. (2007) Evolutionary explanations for cooperation. *Current Biology* **17**:661–72. [aCKWDD]

Wheeler B. (2009) Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society B: Biological Sciences* **276**(1669):3013–18. [aCKWDD]

Whitehouse H. (2018) Dying for the group: Towards a general theory of extreme self-sacrifice. *Behavioral and Brain Sciences* **41**:e192. Available at: https://doi.org/10.1017/S0140525X18000249. [AM]

Whitehouse H. & Lanman J. A. (2014) The ties that bind us: Ritual, fusion, and identification. *Current Anthropology* **55**(6):674–95. [aCKWDD]

Whitehouse H., McQuinn B., Buhrmester M. & Swann W. B. (2014) Brothers in arms: Libyan revolutionaries bond like family. *Proceedings of the National Academy of Sciences USA* **111**(50):17783–85. [aCKWDD, DK]

Wildschut T., Insko C. A. & Pinter B. (2003) The perception of outgroup threat: Content and activation of the outgroup schema. In: *The psychology of group perception: Perceived variability, entitativity, and essentialism*, ed. V. Y. Yzerbyt, C. M. Judd & O. Corneille, pp. 335–59. Psychology Press. [HR]

Willems E. P., Arseneau T. J. M., Schleuning X. & van Schaik C. P. (2015) Communal range defence in primates as a public goods dilemma. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**:20150003. [ANR]

Willems E. P. & Van Schaik C. P. (2017) The social organization of Homo ergaster: Inferences from anti-predator responses in extant primates. *Journal of Human Evolution* **109**:11–21. [aCKWDD]

Williams J.M., Oehlert G.W., Carlis J.V. & Pusey A.E. (2004) Why do male chimpanzees defend a group range? *Animal Behaviour* **68**(3):523–32. [WB]

Wilson M. & Daly M. (1985) Competitiveness, risk taking, and the young male syndrome. *Ethology and Sociobiology* **6**:59–73. [DVB]

Wilson M. L. (2001) *Imbalances of power: How chimpanzees respond to the threat of intergroup aggression*. PhD thesis, Harvard University. [ARR]

Wilson M. L., Boesch C., Fruth B., Furuichi T., Gilby I. C., Hashimoto C., Hobaiter C. L., Hohmann G., Itoh N., Koops K., Lloyd J. N., Matsuzawa T., Mitani J. C., Miungu D. C., Morgan D., Muller M. N., Mundry R., Nakamura M., Pruetz J., Pusey A. E., Riedel J., Sanz C., Schel A. M., Simmons N., Waller M., Watts D. P., White F., Wittig R. M., Zuberbühler K. & Wrangham R. W. (2014) Lethal aggression in Pan is better explained by adaptive strategies than human impacts. *Science* **513**:414–17. [WB]

Wilson M. L. & Glowacki L. (2017) Violent cousins: Chimpanzees, humans, and the roots of war. In: *Chimpanzees and human evolution*, ed. M. Muller, D. Pilbeam & R. Wrangham. Harvard University Press. [WB]

Wilson M. L., Hauser M. D. & Wrangham R. W. (2001) Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour* **61**:1203–16. [ANR]

Wilson M. L., Kahlenberg S. M., Wells M. & Wrangham R. W. (2012) Ecological and social factors affect the occurrence and outcomes of intergroup encounters in chimpanzees. *Animal Behaviour* **83**:277–91. [ANR]

Wilson M.L. & Wrangham R.W. (2003) Intergroup relations in chimpanzees. *Annual Review of Anthropology* **32**:363–92. [ARR]

Wohl M. J. A., Squires E. C. & Caouette J. (2012) We were, we are, will we be? The social psychology of collective angst. *Social and Personality Psychology Compass* **6**:379–91. [PP]

Wood B. M., Watts D. P., Mitani J. C. & Langergraber K. E. (2017) Favorable ecological circumstances promote life expectancy in chimpanzees similar to that of human hunter-gatherers. *Journal of Human Evolution* **105**:41–56. [WB]

Woodworth M. & Porter S. (2002) In cold blood: Characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology* **111**(3):436–45. Available at: https://doi.org/10.1037/0021-843X.111.3.436. [TOP]

Wrangham R. W. (1999) Evolution of coalitionary killing. *Yearbook of Physical Anthropology* **42**:1–30. [WB, ACL, ARR]

Wrangham R. W. (2018) Two types of aggression in human evolution. *Proceedings of the National Academy of Sciences USA* **115**:245–53. Available at: https://doi.org/10.1073/pnas.1713611115. [aCKWDD, ACL]

Wrangham R. W. & Glowacki L. (2012) Intergroup aggression in chimpanzees and war in nomadic hunter-gatherers: Evaluating the chimpanzee model. *Human Nature* **23**:5–29. Available at: https://doi.org/10.1007/s12110-012-9132-1. [WB, ACL, ARR]

Wright T. M. (2014) Territorial revision and state repression. *Journal of Peace Research* **51**:375–87. [aCKWDD]

Yamagishi T. (1986) The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* **51**:110–16. [aCKWDD]

Yamagishi T. (1988) Exit from the group as an individualistic solution to the free-rider problem in the United States and Japan. *Journal of Experimental Social Psychology* **24**:530–42. [rCKWDD]

Yamagishi T. & Kiyonari T. (2000) The group as the container of generalized reciprocity. *Social Psychology Quarterly* **63**:116–32. [rCKWDD]

Yamagishi T. & Mifune N. (2008) Does shared group membership promote altruism? Fear, greed, and reputation. *Rationality and Society* **20**(1):5–30. [NM]

Yamagishi T. & Mifune N. (2016) Parochial altruism: Does it explain modern human group psychology? *Current Opinion in Psychology* **7**:39–43. [NM]

Yamagishi T., Mifune N., Li Y., Shinada M., Hashimoto H., Horita Y., Miura A., Inukai K., Tanida S., Kiyonari T., Takagishi H. & Simunovic D. (2013) Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes* **120**(2): 260–71. [NM]

Yamasue H., Yee J.R., Hurlemann R., Rilling J.K., Chen F.S., Meyer-Lindenberg A. & Tost H. (2012) Integrative approaches utilizing oxytocin to enhance prosocial behavior: From animal and human social behavior to autistic social dysfunction. *Journal of Neuroscience* **32**:14109–17. Available at: http://www.jneurosci.org/content/32/41/14109.long. [RH]

Young I. F. & Sullivan D. (2016) Competitive victimhood: A review of the theoretical and empirical literature. *Current Opinion in Psychology* **11**:30–34. Available at: https://doi.org/10.1016/j.copsyc.2016.04.004. [HR]

Young L. J. & Wang Z. (2004) The neurobiology of pair bonding. *Nature Neuroscience* **7** (10):1048–54. Available at: https://www.nature.com/articles/nn1327. [RH]

Yuki M. & Yokota K. (2009) The primal warrior: Outgroup threat priming enhances intergroup discrimination in men but not women. *Journal of Experimental Social Psychology* **45**:271–74. [DVB]

Zartman I. W. (1989) *Ripe for resolution: Conflict resolution in Africa*. Oxford University Press. [rCKWDD]

Zartman I. W. (2000) Ripeness: The hurting stalemate and beyond. In: *Conflict resolution after the Cold War*, ed. P. C. Stern & D. Druckman, pp. 225–50. National Academy Press. [rCKWDD]

Zerres A. & Hüffmeier J. (2011) Too many cooks spoil the soup. *Die Betriebswirtschaft* **71**:559–75. [JH]

Zhang H., Gross J., De Dreu C. K. W. & Ma Y. (2019) Oxytocin promotes coordinated out-group attack during intergroup conflict in humans. *eLife* **8**; e40698. doi: 10.7554/eLife.40698. [arCKWDD]