UvA-DARE (Digital Academic Repository)

Neurocognitive Underpinnings of Aggressive Predation in Economic Contests

Rojek-Giffin, M.; Lebreton, M.; Scholte, H.S.; van Winden, F.; Ridderinkhof, K.R.; De Dreu, C.K.W.

Link to publication

Neurocognitive Underpinnings of Aggressive Predation in Economic Contests

Michael Rojek-Giffin[1], Mael Lebreton[2], H. Steven Scholte[3], Frans van Winden[3], K. Richard Ridderinkhof[3], and Carsten K. W. De Dreu[1,3]

Abstract

■ Competitions are part and parcel of daily life and require people to invest time and energy to gain advantage over others and to avoid (the risk of) falling behind. Whereas the behavioral mechanisms underlying competition are well documented, its neurocognitive underpinnings remain poorly understood. We addressed this using neuroimaging and computational modeling of individual investment decisions aimed at exploiting one's counterpart ("attack") or at protecting against exploitation by one's counterpart ("defense"). Analyses revealed that during attack relative to defense (i) individuals invest less and are less successful; (ii) computations of expected reward are strategically more sophisticated (reasoning level $k = 4$ vs. $k = 3$ during defense); (iii) ventral striatum activity tracks reward prediction errors; (iv) risk prediction errors were not correlated with neural activity in either ROI or whole-brain analyses; and (v) successful exploitation correlated with neural activity in the bilateral ventral striatum, left OFC, left anterior insula, left TPJ, and lateral occipital cortex. We conclude that, in economic contests, coming out ahead (vs. not falling behind) involves sophisticated strategic reasoning that engages both reward and value computation areas and areas associated with theory of mind. ■

INTRODUCTION

In his *Principles of Political Economy*, John Stuart Mill (1859) observed that "a great proportion of all efforts… [are] spent by mankind in injuring one another, or in protecting against injury." Such appetite for "injuring others" and to defend against being injured has recently been documented in economic contest experiments in which individuals invest to obtain a reward at a cost to their competitor (henceforth attack) or to avoid losing their resources to their antagonist (henceforth defense; De Dreu & Gross, 2019; Chowdhury, Jeon, & Ramalingam, 2018; De Dreu, Kret, & Sligte, 2016; Wittmann et al., 2016; Chen & Bao, 2015; De Dreu, Scholte, van Winden, & Ridderinkhof, 2015; Zhu, Mathewson, & Hsu, 2012; Carter & Anderton, 2001; Grossman & Kim, 1996). These experiments showed that humans invest in injuring others through attacks and in protecting against injuring through defense, that investments in attack are typically less frequent and forceful than investments in defense, and that attack decisions disproportionally often fail and defenders relatively often survive (with ≈30% victories against ≈70% survivals; for a review, see, e.g., De Dreu & Gross, 2019).

Resonating with the idea that competition can be costly, participants during such attacker–defender contests typically waste about 40% of their wealth in fighting each other (De Dreu & Gross, 2019). Yet why people invest in attack and defense remains poorly understood. In fact, investing in injuring others and in protecting against injury may reflect an array of subjective "desires" (Charpentier, Aylward, Roiser, & Robinson, 2017; Delgado, Schotter, Ozbay, & Phelps, 2008; Dorris & Glimcher, 2004). Perhaps humans invest in attack and defense to maximize their personal earnings, as is typically assumed in standard economic theory (e.g., Ostrom, 1998). Relatedly, individuals may invest in attack and defense because of "competitive arousal" and rivalry (Delgado et al., 2008; Ku, Malhotra, & Murnighan, 2005). Finally, investment in attack and defense may be driven by a desire to minimize risk and uncertainty (Delgado et al., 2008; Kahneman & Tversky, 1984). Indeed, decision-making in competitive contests is inherently risky—investments are typically wasted and may result in no return (among attackers), wasted resources (when attacks were unexpectedly shallow and one thus overinvested in defense), or costly defeat (when attacks were unexpectedly tough). Humans factor in such risks when making decisions and are typically risk-averse (Tobler, O'Doherty, Dolan, & Schultz, 2007; Kuhnen & Knutson, 2005; Loewenstein, Weber, Hsee, & Welch, 2001).

Humans may hold conflicting desires when investing in attack and defense and may need to balance between maximizing reward and minimizing risk. What individuals aim for and how possibly conflicting desires are regulated is difficult to infer from behavioral decision-making alone. To illustrate, consider a two-player contest in which one participant can invest in attack and the other participant

[1]Leiden University, [2]University of Geneva, [3]University of Amsterdam

in defense. When the attacker invests more than its defender, attackers obtain all what the defender did not invest, and the defender would be left with 0. If attackers invest equal or less than their defender, both sides earn their noninvested resources (De Dreu & Gross, 2019; Chowdhury et al., 2018; De Dreu, Gross, et al., 2016; De Dreu, Kret, et al., 2016; De Dreu et al., 2015; Carter & Anderton, 2001; Grossman & Kim, 1996).[1] It follows that investments can increase attacker earnings and their competitive success and can prevent defenders from losing their remaining endowment to their attacker. At the same time, however, not investing resources eliminates the attacker's uncertainty about earnings from the contest, alongside the possibility of losing money. Defenders, in contrast, reduce such uncertainty and possibility of losing the contest by investing resources (Chowdhury et al., 2018).

We solved this problem of inference using a two-pronged approach inspired by recent work in cognitive neuroscience on learning from reward and risk prediction (Olsson, FeldmanHall, Haaker, & Hensler, 2018; Palminteri, Wyart, & Koechlin, 2017; Preuschoff, Quartz, & Bossaerts, 2008; Preuschoff & Bossaerts, 2007). First, from investments in attacker–defender contests, we computed, using a k-level reasoning approach, estimates of expected reward and expected risk (Zhu et al., 2012; Ribas-Fernandes et al., 2011; Botvinick, Niv, & Barto, 2009; Camerer, Ho, & Chong, 2004; Nagel, 1995; Stahl & Wilson, 1995; Harsanyi, 1967). The computational approach incorporates the intuition that the formation of expectations and beliefs in strategic interactions are recursive (i.e., [1] I think that [2] you think that [3] I think that [4]…) and can be more or less sophisticated (i.e., the number of recursions k). Using computational modeling and model comparison, we estimated for each investment in attack and defense the expected reward and risk, and concomitant reward and risk prediction errors. Our modeling thus defines (expected) reward as the (expected) monetary payoff from investment in attack and defense (e.g., Zhu et al., 2012) and (expected) risk as the (expected) variance of the reward prediction error (Preuschoff et al., 2008; Preuschoff & Bossaerts, 2007).

Second, and next to an exploratory whole-brain analysis potentially revealing currently unknown cues about the neural foundations of exploitation and protection, we linked prediction errors to a priori defined ROIs—the ventral striatum (VS) and the amygdala. We chose the VS because it has been extensively linked to reward processing and competitive success (viz. reward maximization; Metereau & Dreher, 2015; McNamee, Rangel, & O'Doherty, 2013; Balodis et al., 2012; Rudorf, Preuschoff, & Weber, 2012; Zhu et al., 2012; Xue et al., 2009; Preuschoff & Bossaerts, 2007). We chose the amygdala because of its involvement in low-level affective processing of threat to resources (viz. risk minimization; De Dreu et al., 2015; Choi & Kim, 2010; Baumgartner, Heinrichs, Vonlanthen, Fischbacher, & Fehr, 2008; Delgado et al., 2008; Nelson & Trainor, 2007; Phelps & LeDoux, 2005).

METHODS

Participants and Ethics

Male participants ($M = 25.31$ years, $n = 27$) were recruited via an online recruiting system for participating in a neuroimaging study on human decision-making. Exclusion criteria were significant neurological or psychiatric history, prescription-based medication, smoking more than five cigarettes per day, and drug or alcohol abuse.[2] Eligible participants were assigned to a session and instructed to refrain from smoking or drinking (except water) for 2 hr before the experiment that lasted approximately 1.5 hr. They received a show-up fee of €30 in addition to the earnings from decision-making. The experiment involved no deception and was incentivized (see below), received ethics approval from the Psychology Ethics Committee of the University of Amsterdam, and complied with the guidelines from the American Psychological Association (6th edition). Participants provided written informed consent before the experiment and received a full debriefing afterward.

Experimental Procedures

Experimental sessions were conducted between noon and 4 p.m., and participants were tested individually (also see De Dreu et al., 2015). Upon arrival, participants were escorted to a private cubicle where they read and signed an informed consent form. Participants received a booklet with instructions for the attacker–defender game (labeled Investment Task), containing several examples of investments and their consequences to both attacker (labeled Role A) and defender (labeled Role B) and several questions to probe understanding of the game structure and decision consequences. Neutral labeling was used throughout.

Upon finishing the instructions for the contest, the experimenter prepared the participant for neuroimaging. During the fMRI session, participants completed six functional runs, each consisting of a 20-trial block played as either attacker or defender. Participants thus alternated between the role of attacker and defender every 20 trials, with the starting order counterbalanced across participants. Importantly, we used a random partner matching one-shot protocol, eliminating reputation concerns (Zhu et al., 2012). In each session, participants made 60 investments as attacker and 60 as defender. For each investment trial, they received a prompt, randomly generated between 0 (indicating no investment) and 10 (indicating investment of the entire endowment) and used a button-press to adjust the given number up or down to indicate their desired investment. The duration of the selection period was self-paced and had an average length of 4.27 sec ($SD = 3.43$ sec; see Figure 1). After selecting their investments, participants waited an average of 6.08 sec ($SD = 2.22$ sec), at which point they received feedback about their counterpart's investment and were shown the
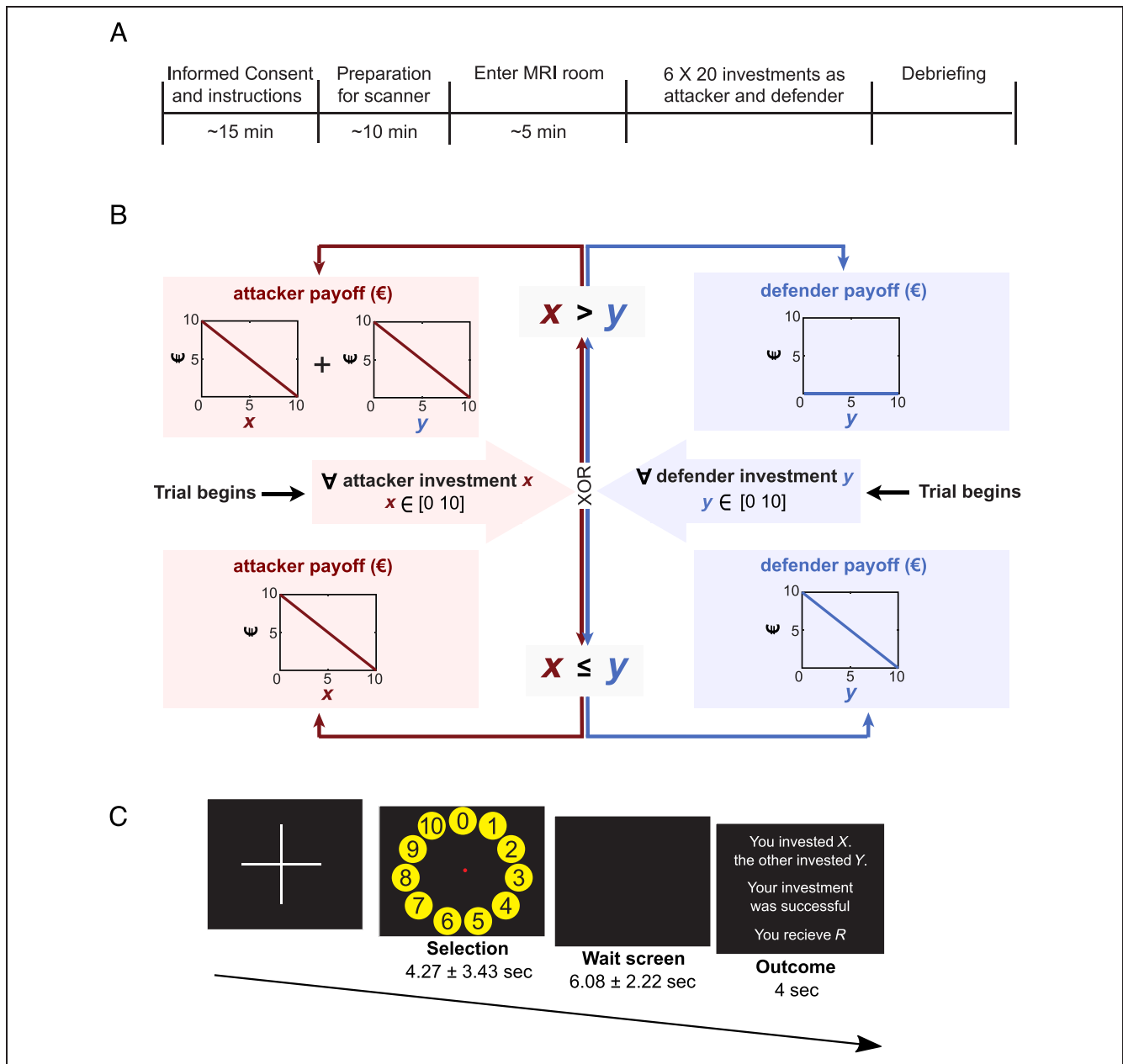
Figure 1. Experimental design. (A) Timeline of the entire experiment. (B) The attacker–defender contest: On each trial, both attacker and defender begin with a €10 endowment with which to invest in the contest. Investments are nonrecoverable, yet if the defender invests equal or more than the attacker (bottom), both attacker and defender keep their remaining endowments (i.e., whatever they did not invest in the contest). If the attacker invests more than the defender (top), the attacker receives their remaining endowment plus that of the defender, who receives nothing. (C) Trial breakdown: For each trial, participants received a prompt, randomly generated between 0 (indicating no investment) and 10 (indicating investment of the entire endowment) and used a button-press to adjust the given number up or down to indicate their desired investment. The duration of the selection period was self-paced ($M = 4.27$ sec, $SD = 3.43$). After selecting their investments, participants waited an average of $M = 6.08$ sec ($SD = 2.22$) and then received feedback about their counterpart's investment and the payoffs to oneself and to the counterpart. This completed one trial.

respective payoffs to oneself and the other (who was randomly chosen on each trial from a pool of 150 attacker [defender] investments; for further detail, see De Dreu, Giacomantonio, Giffin, & Vecchiato, 2019; De Dreu et al., 2015). At the end of the experiment, participants received their participation fee and earnings by bank transfer (range €0–€8, with M = €5 for nonscanner participants, and €0–€33, with M = €19 for scanner participants).

Accordingly, participant pay was private and conditioned on their performance.

Attacker–Defender Contest

The attacker–defender contest (Figure 1B) consists of two players: an attacker and a defender. Each player was endowed with €10 from which they could invest

in the contest. Investments were always wasted, but if the investments by the attacker (x) exceeded that by the defender (y), the attacker ($x > y$) obtains all of the defender's noninvested endowment ($e - y$). In this case, the attacker's total earning was $2e - x - y$, and the defender earned 0. If, in contrast, the defenders investment matched or exceeded that by the attacker ($y \geq x$), both defender and attacker earned what was left from their endowment ($e - y$ and $e - x$, respectively; De Dreu et al., 2015, 2019; De Dreu, Gross, et al., 2016; De Dreu, Kret, et al., 2016).

The attacker–defender contest has a contest success function $f = X^m/(X^m + Y^m)$, where f is the probability that the attacker wins, $m \to \infty$ for $X \neq Y$ and $f = 0$ if $Y = X$. Assuming rational selfish play and risk neutrality, standard economic theory predicts that attackers and defenders use mixed strategies when investing. With $e = €10$ per trial (as used in the current experiment), the mixed strategies for attack (with probability of investing x denoted by $p(x)$) and defense (with probability of investing y denoted by $p(y)$) define a unique Nash equilibrium where expected investments in attack are both lower ($x = 2.62$) than in defense ($y = 3.38$) and less frequent (probability of attack [defense] = 60% [90%]). However, when attacks are made, they are expected to be more "forceful" (4.36 vs. 3.75 for defense).[3]

Modeling Investment Behavior with *k*-level Sophistication

To compute individual estimates of expected reward and concomitant reward and risk prediction errors, we adapted the cognitive hierarchies framework developed in behavioral economics (Botvinick et al., 2009; Camerer et al., 2004; Nagel, 1995). The idea is that players hierarchically form beliefs about their opponents' behavior up to a certain level of cognitive sophistication (*k*-level). A *k*-0 player invests randomly. At $k = 1$, the individual assumes that her opponent has $k = 0$ and finds an investment that maximizes her expected reward under this assumption. At $k = 2$, the individual assumes that her opponent has $k = 1$ and finds an investment that maximizes her own expected reward under the assumption that the opponent seeks to maximize his personal reward against a *k*-0 player. This recursion can, in theory, continue infinitely, yet in our computational modeling, we limited $k \leq 5$. Specifically, when I_s represents a player's own investment (s stands for "self") and I_o as their representation of the other player's investment (o stands for "other"), we can formally express:

k-level 0

k-level 0 play each strategy with equal probability. We have

$$\forall b \in \{0, \ldots, 10\}, P(I_s = b) = \frac{1}{11} \qquad (1)$$

k-level 1

k-level 1 expect their opponent to play as *k*-level 0, such that they expect

$$\forall b \in \{0, \ldots, 10\}, P(I_o = b) = \frac{1}{11} \qquad (2)$$

These expectations can be used to compute the probability of success S of a given investment b ($P(S|b)$) by the attacker A and defender D, respectively

$$\forall b_A \in \{0, \ldots, 10\}, P(S|b_A) = \sum_{i=0}^{b_A - 1} P(I_o = i) \qquad (3)$$

$$\forall b_D \in \{0, \ldots, 10\}, P(S|b_D) = \sum_{i=0}^{b_D} P(I_o = i) \qquad (4)$$

This can be used to compute an expected value, which in this case in the expected reward *ER* for any potential investment by the attacker and defender. We have, for the attacker

$$ER_A(b_A) = [P(S|b_A) \times (10 - \mathbb{E}[b_D|b_D < b_A] + 10 - b_A)] + [(1 - (S|b_A)) \times (10 - b_A)] \qquad (5)$$

where the two square brackets represent cases where the investment is successful or unsuccessful, respectively, and $\mathbb{E}[b_D|b_D < b_A]$ is the expected opponent's investment in case of success

$$\mathbb{E}[b_D|b_D < b_A] = \sum_{i=0}^{b_A - 1} i \times P(I_o = i) \qquad (6)$$

For the defender we have, likewise

$$ER_D(b_D) = [P(S|b_D) \times (10 - b_{At})] + [(1 - P(S|b_D)) \times 0] \qquad (7)$$

The expected reward also has an associated prediction error *PE*, which is simply the expected reward *ER* subtracted from the actual reward *R*

$$PE = R - ER \qquad (8)$$

These values also allow for the calculation of risk prediction *RP* and accompanying risk prediction errors PE_{Risk}. We defined risk prediction as the expected size-squared of the reward prediction error (Preuschoff et al., 2008; Preuschoff & Bossaerts, 2007). More specifically, risk prediction is defined as the sum across all the possible rewards (R) of $(R - ER)^2$ multiplied by the probability $P(R)$ that R is obtained. More formally,

$$RP = \mathbb{E}[(R - ER)^2] = \sum_R P(R) \times (R - ER)^2 \qquad (9)$$

which means that the risk prediction error PE_{Risk} is the risk prediction RP subtracted from the actual size-squared of the reward prediction error

$$PE_{Risk} = (R - ER)^2 - RP \qquad (10)$$

Following standard practices in the field, we assume that participants select the investment I_s that (soft)maximizes their expected reward. This is modeled with a multinomial softmax function with free parameter β, which indexes the exploration/exploitation trade-off (choice temperature),

$$P(I_s = b_i) = \frac{\exp(\beta_1 \times EV(b_i))}{\sum_{j=1}^{10} \exp(\beta_1 \times EV(x_j))} \qquad (11)$$

This choice temperature defines the likelihood of investments I_s, that is, the probability of observing investment I_s under the considered model and parameter values.

k-levels $2 \rightarrow n$

For each k-level, $k \geq 2$, the above procedure is iterated k-times, with k-level predictions of investments—needed to compute probabilities of success, expected rewards, and choice probabilities—being generated by the softmax at the preceding level (see Figure 2). Hence, each k-level model has k-free parameters, which constitutes the choice temperature at each level β_k.

Model Fitting

For each model M, the parameters θ_M ($\theta_M = \{\beta_1, \beta_2, \ldots \beta_k\}$) were optimized by minimizing the negative logarithm of the posterior probability (LPP) over the free parameters

$$LPP = -\log(P(\theta_M | D, M)) \propto -\log(P(D | M, \theta_M)) \qquad (12)$$
$$- \log(P(\theta_M | M))$$

Here, $P(D | M, \theta_M)$ is the likelihood of the data D (i.e., the observed choice) given the considered model M and

Figure 2. Computational framework. Players hierarchically form beliefs about their opponents' behavior, up to a certain level of cognitive sophistication (k-level; Column 1). The expected frequencies of the opponents investment are then used to calculated expected probability of success for each investment (Column 2), which can then be used to calculated expected reward (Column 3). Based on the expected reward, we calculate the frequency that a player should make each investment (Column 4). A k-2 player (Row 2) will assume that her opponent is $k - 1$ and adjust her behavior accordingly, and so on. We developed computational models for hierarchies 1 up to 5.
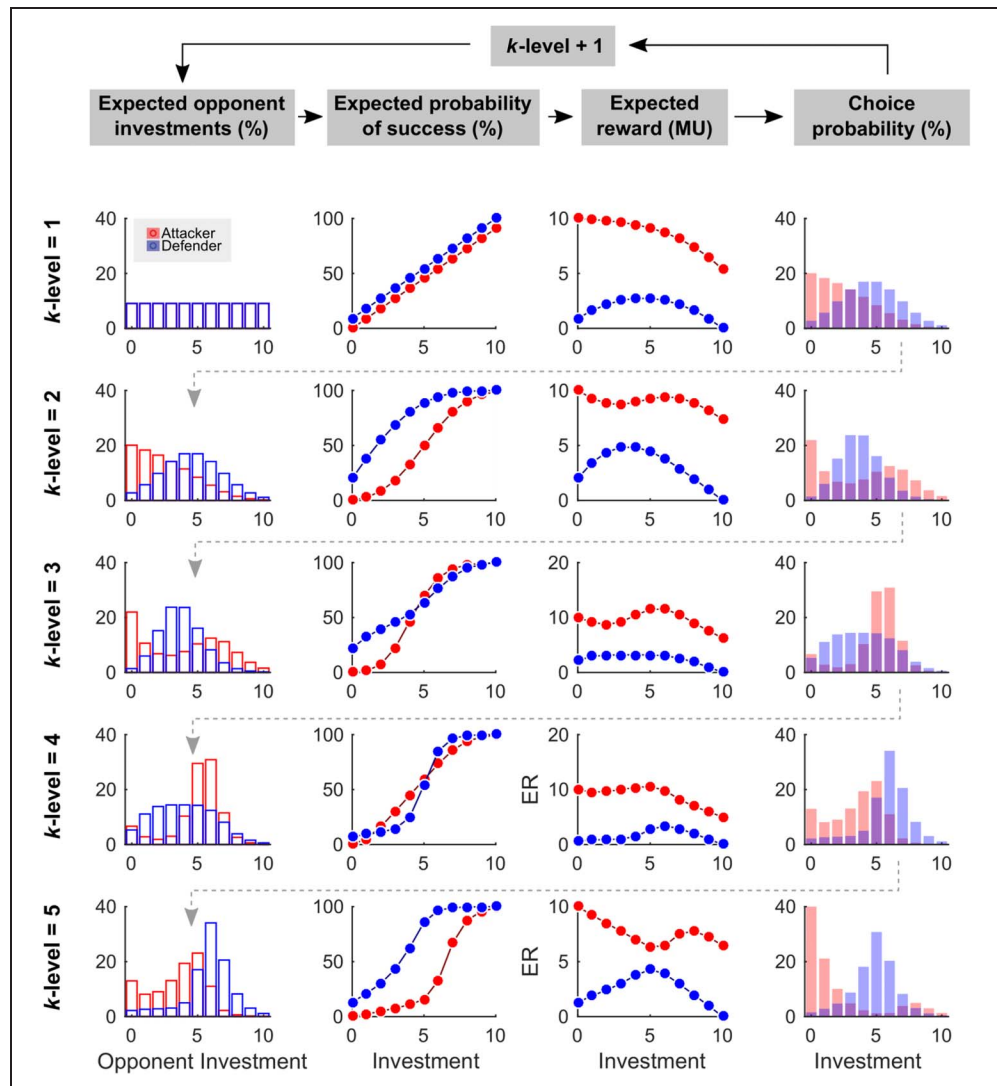
parameter values θ_M, and $P(\theta_M|M)$ is the prior probability of the parameters. Following Daw (2011), the prior probability distributions were defined as a gamma distribution (gampdf(β, 1.2, 5)) for the choice temperature. This procedure was conducted using MATLAB's (The MathWorks, Inc.) *fmincon* function with different initialized starting points of the parameter space (i.e., $0 < \beta <$ Infinite; Palminteri, Khamassi, Joffily, & Coricelli, 2015). We computed the Laplace approximation to model evidence (ME). It measures the ability of each model to explain the experimental data by trading off their goodness-of-fit and complexity. Defining θ_M as the model parameters identified in the optimization procedure and n as the number of data points (i.e., trials), ME was computed as follows (where $|H|$ is the determinant of the Hessian matrix)

$$\text{ME} = \log(P(D|M, \theta_M)) + \log(P(\theta_M|M)) + \frac{df}{2}\log(2\pi) - \frac{1}{2}\log|H| \quad (13)$$

Bayesian Model Comparison

To identify the model most likely to have generated a certain data set, ME was computed at the individual level for each model in the respective model space and fed to random effects Bayesian model comparison using the mbb-vb-toolbox (mbb-team.github.io/VBA-toolbox/; Daunizeau, Adam, & Rigoux, 2014). This procedure estimates the expected frequencies (denoted PP) and the exceedance probability (denoted XP) for each model within a set of models, given the data gathered from all participants. PP quantifies the posterior probability that the model generated the data for any randomly selected participant. XP quantifies the belief that the model is more likely than all the other models of the model space. An XP > 95% for one model within a set is typically considered as significant evidence in favor of this model being the most likely.

Model Identifiability

To assess the reliability of our modeling approach, we performed model identifiability simulations (see Correa et al., 2018, for a similar approach). Choices from synthetic participants were generated for each task and each model by running our computational models, with model parameters sampled in their prior distribution: softmax temperature β were drawn from gamma distribution (random("Gamma," 1.2, 3)). For each model, we ran 10 simulations including 27 synthetic participants ($n = 270$), playing both attacker and defender for three blocks of 20 trials. Model identifiability was assessed by running the Bayesian model comparison on the synthetic data.

MRI Data Acquisition, Preprocessing, and Data Analysis

Scanning was performed on a 3T Philips Achieva TX MRI scanner using a 32-channel head coil. Each participant played six blocks of the attacker–defender game in which functional data were acquired using a gradient-echo, echo-planar pulse sequence (repetition time = 2000 msec, echo time = 27.63 msec, flip angle = 76.18, 280 volumes, field of view = 192^2 mm, matrix size = 64^2, 38 ascending slices, slice thickness = 3 mm, slice gap = 0.3 mm) covering the whole brain. For each participant, we also recorded a 3DT1 recording (3D T1 TFE, repetition time = 8.2 msec, echo time = 3.8 msec, flip angle = 88, field of view = 256^2 mm, matrix size = 256^2, 160 slices, slice thickness = 1 mm) as well as respiration, pulse oximetry signal, and breath rate. Stimuli were back-projected onto a screen that was viewed through a mirror attached to the head coil.

Analyses were conducted with FSL (Oxford Centre for Functional MRI of the Brain Software Library; www.fmrib.ox.ac.uk/fsl) and custom scripts written in MATLAB. All fMRI data were prewhitened, slice-time corrected, spatially smoothed with a 5-mm FWHM gaussian kernel, motion corrected, and high-pass filtered. Functional images were registered to each participant's high-resolution T1 scan and subsequently registered to Montreal Neurological Institute (MNI) space.

Our primary goal was to determine if neural activity was modulated by the expected values and/or prediction errors from our reinforcement learning model. The entire fMRI analysis consisted of a three-level analysis: Level 1 was averaging within runs within participants, Level 2 was averaging across runs within participants, and Level 3 was testing for significance at the group level. We constructed three different general linear models (GLMs) to test for significant neural differences between attack and defense behavior as well as to see if attack and defense behavior correlated with our variables of interest. GLM-1 was meant to test for simple model-free differences between attacker and defender neural activity and consisted only of the selection and feedback epochs. GLM-2 was meant to determine if neural activity significantly correlated with investment magnitude during the selection time-phase and whether wins/losses significantly correlated with neural activity during feedback. To this end, it consisted of the following regressors: selection, selection modulated by investment (orthogonalized with respect to selection), feedback, and feedback modulated by wins/losses (z-scored and orthogonalized with respect to feedback). GLM-3 was meant to determine whether any neural activity correlated with the parameters calculated from our k-level model and contained the following regressors: selection, selection modulated by expected value (orthogonalized with respect to selection), selection delayed by 4 sec to capture the delayed nature of risk prediction (Preuschoff et al., 2008), delayed selection modulated by risk prediction (orthogonalized

with respect to delayed selection), feedback, feedback modulated by the prediction error (z-scored and orthogonalized with respect to feedback), and feedback modulated by the risk prediction error (z-scored orthogonalized with respect to feedback). To mitigate spurious results from asymmetric parameter value ranges (Lebreton, Bavard, Daunizeau, & Palminteri, 2019), each parametric regressor was z-scored within each role, meaning both attacker and defender parametric regressors had identical variance.

We checked for multicollinearity by calculating the variance inflation factors (VIFs) for each regressor of interest (Mumford, Poline, & Poldrack, 2015) and found none to be problematic (all VIFs < 2.3). However, four participants made identical investments on every trial, which resulted in rank deficient models (four participants for GLM-2 and GLM-3). Specifically, two individuals made the exact same investment on all attack decisions, one individual made the exact same investment on all defense decisions, and one individual made the exact same investment during attack and defense. These participants had to be removed from the analysis. We tested for an interaction effect between role and each variable of interest by contrasting the relevant parameter estimates for attack and defense in a second-level within-participant fixed-effects analysis. Finally, we tested for group-level significance and corrected for multiple comparisons using FSL's FLAME 1 with the standard cluster forming threshold of $Z > 3.1$ and clusters significant at $p = .05$. We ran additional control analyses with FSL's randomized threshold-free cluster enhancement (Winkler, Ridgway, Webster, Smith, & Nichols, 2014; Smith & Nichols, 2009), and results were virtually identical.

We also conducted analyses within an a priori selected anatomical VS and within an a priori selected anatomical amygdala ROI. Both masks were obtained from the meta-analytic tool Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). We used the terms "ventral striatum" and "amygdala" in our search of Neurosynth, instead of using "reward" or "fear." Avoiding psychological constructs such as reward or fear reduced possible bias in our ROIs in favor of a particular psychological construct. For our ROI analyses, we took the average value across every voxel within each ROI for each participant within the contrast of interest (e.g., attacker–reward prediction error) and then tested for significance with a paired-sample t test.
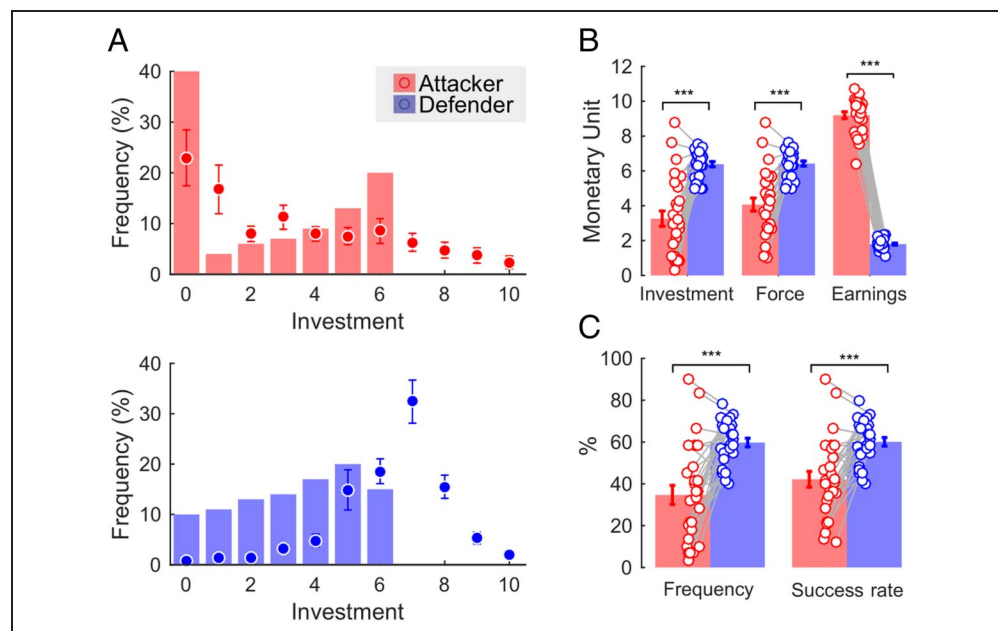
RESULTS

All codes used to produce analyses and plots can be found at 10.6084/m9.figshare.11877699. And all unthresholded statistical brain maps can be found at https://neurovault.org/collections/6740/.

Decision-making

Earlier reports of the attacker–defender contest game analyzed investments in terms of the overall investment (range 0–10), the frequency of investment (all trials in which x or $y > 0$; range 0–60), and the force of investment (the amount invested on nonzero investment trials; range 1–10). For these measures, we find, consistent with earlier work, that individuals invested less often in attack than in defense, $t(26) = -4.12$, $p = .0003$; invested in attack less overall, $t(26) = -8.56$, $p < .0001$; and invested less forcefully in attack than in defense, $t(26) = -7.81$, $p < .0001$ (Figure 3B). Although individuals earned more from attack (noninvested resources + spoils of winning) than defense trials (noninvested resources in case of

Figure 3. Behavioral results.
(A) Nash equilibrium predictions (bars) plotted against empirical distribution of participants' investments (dots with error bars are means ± 1 *SE*) for attacker (top row, red) and defenders (bottom row, blue).
(B) Attacker (red) and defender (blue) investments, force of investment, and mean earnings (shown are means ± 1 *SE*).
(C) Frequency of investment and success rate (shown are means ± 1 *SE*). Contrasts are significant at *p < .05, **p < .01, and ***p < .001.

Figure 4. Brain imaging results. Whole-brain analysis testing for attacker neural activity correlated to wins and losses (A) and feedback differences between attacker and defender (B). (A) Wins and losses as an attacker correlated with neural activity in the TPJ, IFG, VS, anterior insula (AI), thalamus (THA), and lateral occipital cortex (LOC). (B) Processing feedback as an attacker associated with more neural activation in the left IFG, left AI, and left OFC. All contrasts are FWE-corrected at $p < .05$ for the whole brain.
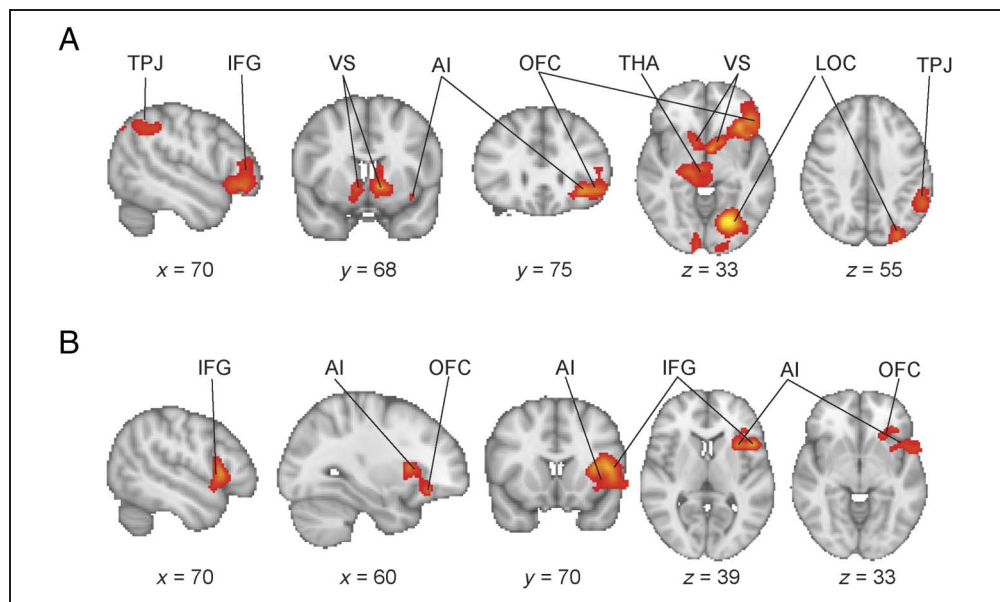
survival), $t(26) = 43.91, p < .0001$, they were less successful during attack than defense trials, $t(26) = -7.22, p < .0001$: As defender they "survived" more often than that they "killed" as attacker (Figure 3C).

In addition to the contrast between attack and defense, we examined investments in relation to predictions derived from standard economic theory that assumes rational self-interest and risk neutrality. Relative to mixed-strategy equilibrium predictions (see Methods), individuals invest more and more forcefully in defense ($t(26) = 20.40$, $p < .0001$, and $t(26) = 18.467, p < .0001$, respectively), but not more and not more forcefully in attack ($t(26) = 1.46, p = .157$, and $t(26) = -0.78, p = .441$, respectively; Figure 3A). Still, however, both attack and defense returned less earnings than predicted by standard economic theory ($t(26) = -4.19, p = .00028$, and $t(26) = -40.56, p < .0001$), and the frequency of both attacks and defense exceeded expectations based on rational selfish play ($t(26) = 3.04, p = .0054$, and $t(26) = 30.26, p < .0001$, respectively). Conversely, success rates for attacks (victories) and defense (survival) did not deviate from Nash equilibrium predictions ($t(26) = -0.25, p = .804$, and $t(26) = -0.98, p = .336$, respectively).

Neural Correlates of Attack and Defense

To examine the neural foundations of decision-making during attack and defense, we performed whole-brain analyses on the selection phase (when participants decided whether and how much to invest in attack or defense) and on the feedback phase (when participants received information about their opponent's investment and the resulting outcomes to oneself). Whereas no significant differences between attacker and defender were observed during selection, whole-brain analyses did show significant attacker–defender contrasts for the feedback phase. Specifically, during feedback, participants exhibited higher BOLD response during attack relative to defense in a cluster within the left anterior insula and inferior frontal gyrus (IFG; Figure 4: MNI coordinates: $x = -40, y = 10, z = 16, Z = 4.88$, cluster size = 1657, $p = .0151$, family-wise error [FWE]-whole brain).

Table 1. Regions Exhibiting Significant Correlation between Neural Activity and Win/Loss Feedback during Attack

Region	Peak			Cluster Size	z Value	p (FWE-corr)
	x	y	z			
Attacker Win/Loss						
VS/OFC/insula/thalamus	−8	4	−4	5329	4.27	<.001
Lateral occipital cortex	−22	−74	−8	1686	4.75	.002
Occipital pole	8	−84	4	1603	4.45	.002
TPJ/lateral occipital cortex	−26	−84	46	1577	4.1	.003

All statistics are corrected for multiple comparison with FSL's FLAME 1.

In a follow-up analysis, we examined whether participants exhibited a correlation between neural activity and investments (during decision-making) and outcome (win/loss) during feedback. As before, no significant correlations were found between neural activity and investments during attack or defense, nor did the correlation differ between the two roles. During feedback, however, neural activity during attack covaried with wins and losses in clusters that included the bilateral VS, left OFC, left anterior insula, left TPJ, and lateral occipital cortex (Table 1 and Figure 4B). Activity in these same areas also correlated with wins/losses more during attack than defense but did not survive cluster-based multiple comparison correction (with $p < .05$, uncorrected). When participants processed feedback as defenders, there were no clusters that significantly covaried with wins and losses.

Model-based Analyses of Decision-making and Neural Activity

As noted in Methods section, we captured the computations at hand in attack and defense behavior using the cognitive hierarchies framework developed in behavioral economics (Botvinick et al., 2009; Camerer et al., 2004; Nagel, 1995). The idea is that players hierarchically form beliefs about their opponents' behavior, up to a certain level of cognitive sophistication (k-level; see Figure 2). We developed such computational models for hierarchies 1 up to 5 (see Methods section), and first verified that the behavior predicted by different levels of the cognitive hierarchies could be discriminated (see Methods/

Model Identifiability section and Figure 5). We then fitted those models to our participants' investment data and ran a Bayesian model comparison to identify the hierarchy most likely to generate attacker- and defender-like behavior. Our results show that attackers are best described by a model with four levels of recursion (model K4, exceedance probability = 67.20%), whereas defenders are best described by a model with three levels of recursion (model K3, exceedance probability = 87.41%; Figure 5). From these models, we estimated, for each participant and each investment in attack and defense, the expected reward, risk prediction, and concomitant reward and risk prediction errors. These reward and risk prediction errors were then related to neural activity, using both whole-brain and ROI-based analyses.
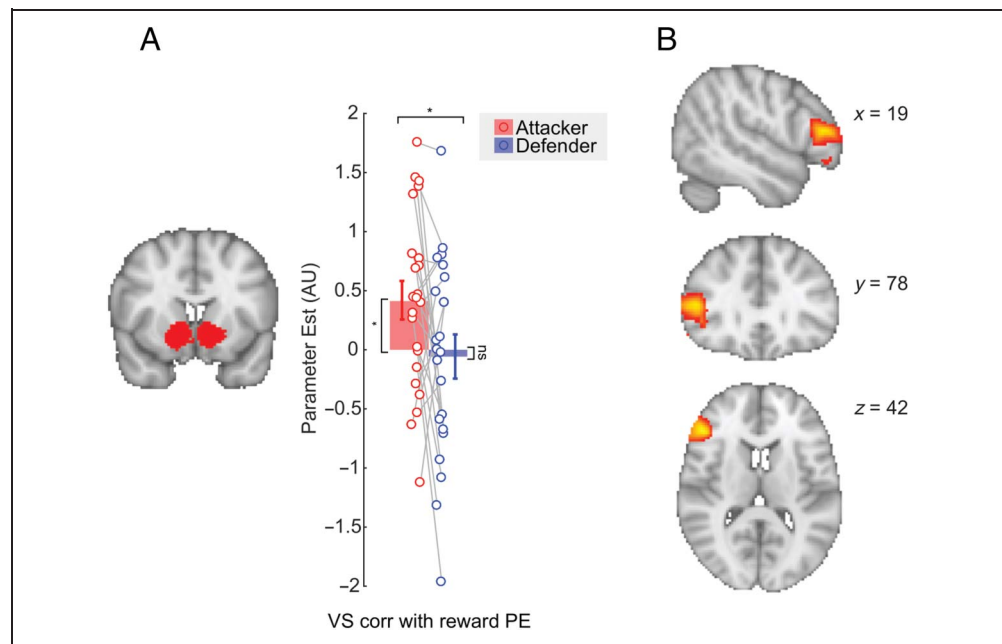
Neural Correlates of Reward Prediction Errors

Within our VS ROI, there was a significant correlation between reward prediction errors and VS neural activity during attack, $t(22) = 2.645$, $p = .0148$, but not during defense, $t(22) = -0.330$, $p = .745$. Furthermore, this correlation between reward prediction errors and VS activity was stronger in attackers than in defenders, $t(22) = 2.189$, $p = .0395$ (see Figure 6A). Within our amygdala ROI, there was no significant correlation between neural activity and reward prediction errors during either attack, $t(22) = 1.785$, $p = .088$, or defense, $t(22) = -1.507$, $p = .146$, but there was a significant difference in correlations between the two roles, $t(22) = 2.405$, $p = .025$.
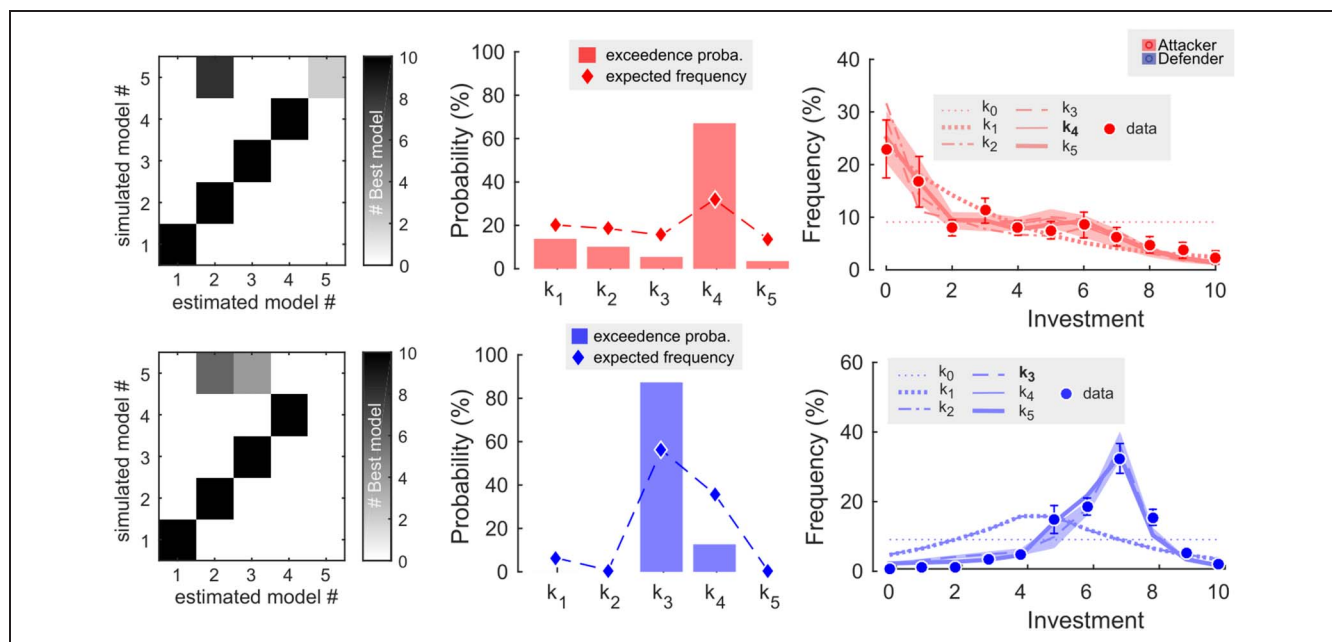
Figure 5. Computational results. (A) Model identifiability, true model used to generate the simulated data (y axis), and the model estimated as most likely based on our Bayesian model comparison (x axis) for both attacker (top row) and defender (bottom row). (B) Exceedance probability (bars) and estimated model frequencies (diamonds) for both attacker (top row) and defenders (bottom row) of each model fit to participant data. (C) Estimates of each model shown in comparison to true behavioral data for both attacker (top row) and defender (bottom row).

Figure 6. Reward prediction errors differentially relate to attacker and defender neural activity. (A) ROI analysis reveals prediction errors during attack significantly correlate with VS activity in attackers but not in defenders. (B) Whole-brain analysis reveals that prediction errors during attack significantly correlate with IFG neural activity. Contrast is FWE-corrected at $p < .05$ for the whole brain.

At the whole-brain level, we found a cluster in the right IFG that significantly correlated with reward prediction errors during attack (MNI coordinates: $x = 48$, $y = 32$, $z = 12$, $Z = 4.55$, cluster size = 681, $p = .0391$, FWE-whole brain; see Figure 6B). We note that this cluster is similar in location to regions found to covary with RTs, but in the present case the correlation between RT and reward prediction errors was not significant ($r = -.0079$, $p = .654$). Because all the contrasts reported were conducted at the feedback time-phase, with the selection time-phase as a covariate RT was at least partially captured by our GLM. Accordingly, because RT reward prediction error is nonsignificant here and RT is captured in the duration of the selection-phase decision-making, we can conclude that RT is not of relevance here.

There were no clusters at the whole-brain level that correlated with reward prediction errors during defense, nor were there any clusters that showed a significant difference in correlation between attacker and defender trials.

Neural Correlates of Risk Prediction Errors

We found that, within our VS ROI, there was no significant correlation between neural activity and risk prediction errors during either attack, $t(22) = -1.622$, $p = .117$, or defense, $t(22) = 0.164$, $p = .871$, nor was there a significant difference in correlations between the two roles, $t(22) = -1.505$, $p = .145$. The same was true in our amygdala ROI (attacker: $t(22) = -0.588$, $p = .562$; defender: $t(22) = 0.363$, $p = .720$; attacker vs. defender: $t(22) = -0.647$, $p = .523$) and at a the whole-brain level.

DISCUSSION

Competition requires that people expend resources to win from other contestants and to expend resources to prevent losing from other contestants. These two core motives operating during competition—coming out ahead versus not falling behind—were examined here in a simple attacker–defender contest in which opposing individuals simultaneously invested, out of a personal endowment, into exploitative attacks and protective defense. As shown by others already, we find here too that individuals invest less frequently and less intensely in economically "injuring others" than they invest in defending themselves against the threat of being economically injured (De Dreu & Gross, 2019, for a review). Computationally, we found that during attack individuals tend to utilize a higher level of cognitive recursion than during defense. We furthermore found attack behavior relative to defense behavior to be preferentially associated with neural regions associated with theory of mind and, within the VS, to be preferentially correlated with reward prediction errors.

What remained poorly understood is why and how people design their strategies of attack and defense. We argued that, in addition to reward maximization, investments in attack and defense may be driven by the desire to out-compete the protagonists as well as by the desire to minimize risk. We approached this issue with a computational framework modeling reward and risk prediction errors based on k-level reasoning in belief formation (Zhu et al., 2012; Camerer et al., 2004; Nagel, 1995). Our results at the neural level revealed no evidence for risk minimization. Instead and in line with earlier work (e.g., Zhu et al., 2012), we find good evidence that contestants aimed to maximize reward both during attack and defense. At the same time, however, we observed significant differences in the computation of expected reward and in the underlying neural activation during attack versus defense. Specifically, we found reward prediction errors during attack (more than during

defense) to robustly correlate with neural activity in the VS and, using whole-brain analyses, the IFG.

Our computational modeling demonstrated that investments in attack are best fitted by a model containing four levels of recursion whereas investments in defense are best fitted by a model containing three levels of recursion. This suggests that individuals engage in more sophisticated reasoning about their protagonist's strategy during attack than defense. Indeed, our neuroimaging results revealed significant attack–defense contrasts in neural activation in regions often associated with perspective taking and "theory of mind"—the lateral occipital cortex, the IFG, and the TPJ (Engelmann, Schmid, De Dreu, Chumbley, & Fehr, 2019; Prochazkova et al., 2018; Van Overwalle, 2009). These results resonate with earlier work showing that temporarily dysregulating the IFG through theta burst stimulation affected investment behavior during attack but not defense (De Dreu, Kret, et al., 2016) and that reducing cognitive capacity before decision-making influenced attackers but not defenders (De Dreu et al., 2019). Combined, these results suggest that individuals engage neural regions for perspective taking and theory of mind during economic contests to out-smart and exploit their protagonist.

Results for neural activity were specific to the feedback phase, when contest outcomes were presented, and not observed during the selection phase, when investment decisions were implemented. Possibly, different neurocognitive operations govern implementation and processing of feedback. During implementation, controlled deliberation may be more or less active, and this may relate to activity in prefrontal regions involved in executive control. Perhaps, the extent to which cognitive control and deliberation during selection is engaged is not conditioned by the specific role decision-makers perform. During feedback, learning and updating operations may be active, and this may relate to neural activation in regions involved in value computation and emotion processing (Behrens, Hunt, & Rushworth, 2009; Yacubian et al., 2006). Indeed, we found neural activity in the VS to be meaningfully related to reward prediction errors (also see Stallen et al., 2018; Zhu et al., 2012; Yacubian et al., 2006; O'Doherty et al., 2004). In contrast to expectations, however, we did not find differential activity in the amygdala, nor amygdala activity to be related to behavioral indicators processed during feedback. Possibly, contestants process feedback in an emotionally detached and rather cognitive manner aimed at revising and updating their (future) strategy for attack and defense.

Our study design included male participants, and extrapolating conclusions to female participants may be nontrivial. Intuitively competitive success and reward maximization may fit an (evolved) male psychology, whereas risk minimization fits an (evolved) female psychology (Niederle & Vesterlund, 2011; Croson & Gneezy, 2009; Spreckelmeyer et al., 2009). At the same time, however, male and female participants tend to perform similarly in the attacker–defender contest studied here (De Dreu et al., 2019). Future work is needed to test whether the neurocognitive mechanisms are similar as well, which would further contradict the intuitive hypothesis derived from evolutionary psychology.

Competitions are part and parcel of human life and can be wasteful. In the current contest, participants destroyed roughly 40% of their wealth in attempts at "injuring others and protecting against being injured" (viz. Mill, 1859). Our neurocomputational approach suggested that injuring others is done through rather sophisticated cognitive reasoning, with the key aim to understand the protagonist's strategy selection such that personal rewards can be optimized. When investing in attack more than in defense, people engage more sophisticated cognitive recursion. Furthermore, neural structures associated with theory of mind and reward processing are recruited more during attack than defense decisions. Perhaps, mentalizing not only serves empathy and prosocial decision-making, but also the strategic goal of reward maximization through exploitation and subordination.

Notes

1. The attack–defense contest belongs to a class of asymmetric conflict games in which one player competes to maximize personal gain and the counterpart competes to prevent exploitation (De Dreu & Gross, 2019; Dechenaux, Kovenock, & Sheremeta, 2015). Including in this class of asymmetric games are the hide-and-seek game (Bar-Hillel, 2015; Flood, 1972), the matching pennies game (Goeree, Holt, & Palfrey, 2003), the inspection game (Nosenzo, Offerman, Sefton, & van der Veen, 2014), and the best shot/weakest link game (Chowdhury & Topolyan, 2016; Clark & Konrad, 2007). Across these games, humans invest to maximize wealth and/or to minimize risk of losing.
2. The sample was the same as used in De Dreu et al. (2015), which used a crossover design to examine the behavioral and

neural effects of oxytocin (vs. placebo) administration. Here, we only analyze investments made under placebo. Moreover, our earlier report only considered trials in which participant decisions affected themselves only and did not include those decision trials in which decisions also affected two other individuals within their group. Here, we include also those previously unanalyzed trials. Because this manipulation revealed no differences, we collapsed across these two conditions. In short, the current study shares 25% of its analyzed data with the previous one, asks a different research question, and uses distinctly different analytic techniques.

3. Specifically, the mixed-strategy equilibrium is computed as follows: Attack: $p(x = 1) = 2/45$, $p(x) = p(x - 1)[(12 - x)/(10 - x)]$ for $2 \leq x \leq 6$, $p(x = 0) = 1 - [p(x = 1) + \ldots + p(x = 6)] = 0.4$, and $p(x) = 0$ for $x \geq 7$; Defense: $p(y) = 1/(10 - y)$ for $0 \leq y \leq 5$, $p(y = 6) = 1 - [p(y = 0) + \ldots + p(y = 5)] = 0.15$, and $p(y) = 0$ for $y \geq 7$ (also see De Dreu et al., 2015).

REFERENCES

Balodis, I. M., Kober, H., Worhunsky, P. D., Stevens, M. C., Pearlson, G. D., & Potenza, M. N. (2012). Diminished frontostriatal activity during processing of monetary rewards and losses in pathological gambling. *Biological Psychiatry*, *71*, 749–757.

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science*, *10*, 419–433.

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*, 639–650.

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science*, *324*, 1160–1164.

Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*, 262–280.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics*, *119*, 861–898.

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

Charpentier, C. J., Aylward, J., Roiser, J. P., & Robinson, O. J. (2017). Enhanced risk aversion, but not loss aversion, in unmedicated pathological anxiety. *Biological Psychiatry*, *81*, 1014–1022.

Chen, S., & Bao, F. S. (2015). Linking body size and energetics with predation strategies: A game theoretic modeling framework. *Ecological Modelling*, *316*, 81–86.

Choi, J.-S., & Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 21773–21777.

Chowdhury, S. M., Jeon, J. Y., & Ramalingam, A. (2018). Property rights and loss aversion in contests. *Economic Inquiry*, *56*, 1492–1511.

Chowdhury, S. M., & Topolyan, I. (2016). The attack-and-defense group contests: Best shot versus weakest link. *Economic Inquiry*, *54*, 548–557.

Clark, D. J., & Konrad, K. A. (2007). Asymmetric conflict: Weakest link against best shot. *Journal of Conflict Resolution*, *51*, 457–469.

Correa, C. M. C., Noorman, S., Jiang, J., Palminteri, S., Cohen, M. X., Lebreton, M., et al. (2018). How the level of reward awareness changes the computational and electrophysiological

signatures of reinforcement learning. *Journal of Neuroscience*, *38*, 10338–10348.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*, 448–474.

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Computational Biology*, *10*, e1003441.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision making, affect, and learning* (Vol. 6, pp. 3–38). Oxford: Oxford University Press.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, *18*, 609–669.

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*, *148*, 1767–1781.

De Dreu, C. K. W., & 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.

De Dreu, C. K. W., Gross, J., Méder, Z., Giffin, M., Prochazkova, E., Krikeb, J., et al. (2016). In-group defense, out-group aggression, and coordination failures in intergroup conflict. *Proceedings of the National Academy of Sciences, U.S.A.*, *113*, 10524–10529.

De Dreu, C. K. W., Kret, M. E., & Sligte, I. G. (2016). Modulating prefrontal control in humans reveals distinct pathways to competitive success and collective waste. *Social Cognitive and Affective Neuroscience*, *11*, 1236–1244.

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*, *10*, 721–728.

Delgado, M. R., Schotter, A., Ozbay, E. Y., & Phelps, E. A. (2008). Understanding overbidding: Using the neural circuitry of reward to design economic auctions. *Science*, *321*, 1849–1852.

Dorris, M. C., & Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*, *44*, 365–378.

Engelmann, J. B., Schmid, B., De Dreu, C. K. W., Chumbley, J., & Fehr, E. (2019). On the psychology and economics of antisocial personality. *Proceedings of the National Academy of Sciences, U.S.A.*, *116*, 12781–12786.

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

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.

Grossman, H. I., & Kim, M. (1996). Predation and accumulation. *Journal of Economic Growth*, *1*, 333–350.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I–III: Part I. The basic model. *Management Science*, *14*, 159–182.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341–350.

Ku, G., Malhotra, D., & Murnighan, J. K. (2005). Towards a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions. *Organizational Behavior and Human Decision Processes*, *96*, 89–103.

Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*, 763–770.

Lebreton, M., Bavard, S., Daunizeau, J., & Palminteri, S. (2019). Assessing inter-individual differences with task-related

functional neuroimaging. *Nature Human Behaviour*, *3*, 897–905.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, *16*, 479–485.

Metereau, E., & Dreher, J.-C. (2015). The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events. *Cortex*, *63*, 42–54.

Mill, J. S. (1859). *On Liberty*. New York: Walter Scott Publishing.

Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *PLoS One*, *10*, e0126255.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, *85*, 1313–1326.

Nelson, R. J., & Trainor, B. C. (2007). Neural mechanisms of aggression. *Nature Reviews Neuroscience*, *8*, 536–546.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, *3*, 601–630.

Nosenzo, D., Offerman, T., Sefton, M., & van der Veen, A. (2014). Encouraging compliance: Bonuses versus fines in inspection games. *Journal of Law, Economics, and Organization*, *30*, 623–648.

O'Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*, 452–454.

Olsson, A., FeldmanHall, O., Haaker, J., & Hensler, T. (2018). Social regulation of survival circuits through learning. *Current Opinion in Behavioral Sciences*, *24*, 161–167.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action: Presidential address, American Political Science Association, 1997. *American Political Science Review*, *92*, 1–22.

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*, 8096.

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, *21*, 425–433.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, *48*, 175–187.

Preuschoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Annals of the New York Academy of Sciences*, *1104*, 135–146.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, *28*, 2745–2752.

Prochazkova, E., Prochazkova, L., Giffin, M. R., Scholte, H. S., De Dreu, C. K. W., & Kret, M. E. (2018). Pupil mimicry promotes trust through the theory-of-mind network. *Proceedings of the National Academy of Sciences, U.S.A.*, *115*, E7265–E7274.

Ribas-Fernandes, J. J. F., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., et al. (2011). A neural signature of hierarchical reinforcement learning. *Neuron*, *71*, 370–379.

Rudorf, S., Preuschoff, K., & Weber, B. (2012). Neural correlates of anticipation risk reflect risk preferences. *Journal of Neuroscience*, *32*, 16683–16692.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, *44*, 83–98.

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., et al. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective Neuroscience*, *4*, 158–165.

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*, 218–254.

Stallen, M., Rossi, F., Heijne, A., Smidts, A., De Dreu, C. K. W., & Sanfey, A. G. (2018). Neurobiological mechanisms of responding to injustice. *Journal of Neuroscience*, *38*, 2944–2954.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*, 1621–1632.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*, 829–858.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage*, *92*, 381–397.

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. S. (2016). Self-other mergence in the frontal cortex during cooperation and competition. *Neuron*, *91*, 482–493.

Xue, G., Lu, Z., Levin, I. P., Weller, J. A., Li, X., & Bechara, A. (2009). Functional dissociations of risk and reward processing in the medial prefrontal cortex. *Cerebral Cortex*, *19*, 1019–1027.

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, *26*, 9530–9537.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*, 665–670.

Zhu, L., Mathewson, K. E., & Hsu, M. (2012). Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, 1419–1424.