Low competitive status elicits aggression: behavioral and neural evidence

Macià Buades-Rotger[1,2,3*], Martin Göttlich[1], Ronja Weiblen[1,4],

Pauline Petereit[1], Thomas Scheidt[1], Brian G. Keevil[5], Ulrike M. Krämer[1,2]

[1] Department of Neurology, University of Lübeck, Lübeck, Germany

[2] Department of Psychology, University of Lübeck, Lübeck, Germany

[3] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

[4] Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany

[5] Department of Clinical Biochemistry, University Hospital of South Manchester, Manchester, UK.

* Corresponding author:

Macià Buades-Rotger, PhD

Radboud University

Montessorilaan 3

6525 HR Nijmegen

The Netherlands

+24 3655935

m.buadesrotger@donders.ru.nl

ORCID: http://orcid.org/0000-0003-1219-9216

Acknowledgements:

This study was funded by the German Science Foundation (grant number KR3691/5-1). We are grateful to Susanne Schellbach and Mourad Zoubir for their help with the data acquisition and to our subjects and confederates for their participation. The authors report no conflicts of interest.

Abstract

Winners are commonly assumed to compete more aggressively than losers. Here, we find overwhelming evidence for the opposite. We first demonstrate that low-ranking teams commit more fouls than they receive in top-tier soccer, ice hockey, and basketball leagues. We replicate this effect in the laboratory, showing that participants deliver louder sound blasts to a rival when placed in a low-status position. Using neuroimaging, we characterize brain activity patterns that encode competitive status as well as those that facilitate status-dependent aggression. These analyses reveal three key findings. First, anterior hippocampus and striatum contain multivariate representations of competitive status. Second, interindividual differences in status-dependent aggression are linked with a sharper status differentiation in the striatum and with greater reactivity to status-enhancing victories in the dorsal anterior cingulate cortex. Third, activity in ventromedial, ventrolateral, and dorsolateral prefrontal cortex is associated with trial-wise increases in status-dependent aggressive behavior. Taken together, our results run counter to narratives glorifying aggression in competitive situations. Rather, we show that those in the lower ranks of skill-based hierarchies are more likely to behave aggressively and identify the potential neural basis of this phenomenon.

Keywords: aggression, status, competition, fMRI, neuroimaging

Introduction

Winners are often taken to be aggressive, not only in sportive environments [1,2] but in society at large [3–6]. Contradicting this widely held assumption, the lower sense of control and the frustration occasioned by frequent defeats is likely to fuel aggression in subordinate competitors [7,8]. Furthermore, accomplished contestants can rely on their ability, whereas less proficient ones might compete more aggressively to outweigh the skill differential. Supporting this idea, observational studies indicate that competitive aggression is most often used to hinder a rival's performance [9,10]. Hence, contrary to popular belief, those in the lower ranks of skill-based hierarchies might be more prone to use aggression than those who win consistently. If that were the case, social narratives emphasizing aggressiveness as a requisite for competitive success would be misguided. Here, we tested whether competitive status is related to physical aggression and investigated the neural underpinnings of this association.

Status-based hierarchies order social life and are hence imprinted in human neurobiology [11]. The rostral aspect of the medial prefrontal cortex (mPFC) has been postulated as a crucial region for inferring a rival's dominance during competitive interactions [12], which might be partially attributable to this region's role in encoding stable behavioral traits [13]. The amygdala and anterior cingulate cortex (ACC), on the other hand, have been suggested to track moment-to-moment changes in status-based hierarchies [14,15]. The ventral striatum (VS) and ventromedial prefrontal cortex (vmPFC) are highly responsive to competitive outcomes and might thus also contribute to acquire status representations during skill-based contests [12,16]. In addition, there is evidence suggesting that the hippocampus encodes relational knowledge of social hierarchies [15,17], along with regions involved in social cognition such as the precuneus and the temporoparietal junction (TPJ) [16,18,19]. Thus, the current neuroscientific evidence permits to formulate a tentative description of how competitive hierarchies are processed in the human brain. First, regions of the *salience network* (amygdala, ACC) [20] capture dominance signals and status changes during social interactions. Second, regions involved in outcome valuation (VS, vmPFC) contribute to hierarchy learning on the basis of observed and/or experienced victories and defeats. Finally, the mentalizing network (precuneus, TPJ, dorsal and rostral mPFC) and the hippocampus enable the encoding and implementation of status hierarchies during competitive decision-making. The key question that we addressed here is which role these neurocognitive processes play in status-dependent aggression.

A potentially crucial neuromodulator of status-seeking behavior is the steroid hormone testosterone. Though generally associated with aggression [21], recent evidence indicates that testosterone can flexibly promote prosocial behavior in order to improve one's social status [22]. For instance, a study showed that endogenous testosterone was positively associated with generosity in low-status (i.e. junior) rugby players, whereas the relationship was *negative* in high-status (i.e. senior) players [23]. Similarly, men with relatively higher testosterone concentrations punished unfair offers more severely, but rewarded fair offers more generously [24]. Given that high-testosterone individuals are generally more sensitive to status cues [25], and that competitive situations inherently invoke the use of aggressive strategies, testosterone might increase competitive aggression by modulating neural reactivity to status signals. The relationship between status-seeking behavior and testosterone might however be moderated by cortisol, a stress-related hormone that is thought to partly suppress testosterone's effects [26]. In the present study, we drew on pre-existing sports data, a behavioral experiment and a neuroimaging study to test a) whether competitive status is linked with aggression, b) whether this is contingent on the neural processing of status signals, and c) whether testosterone amplifies aggressive and/or neural responses to status cues, either by itself or interaction with cortisol.

Results

In a first correlational analysis, we probed whether competitive status was associated with aggressive play across seasons in soccer, basketball and ice hockey. We defined competitive status as a team's position at the end of the regular season, with higher values indicating a lower standing in the rankings. Aggression was operationalized as the number of fouls (football), personal fouls (basketball), or penalty time (ice hockey) incurred relative to received, a measure that accounts for the reciprocal, "tit-for-tat" character of sports aggression (see Methods). We collapsed these variables across a minimum of ten seasons between the years 2000 and 2019 separately for the main five European football leagues (England, Spain, Germany, Italy, and France), the North-American National Basketball Association league (NBA), and the North-American National Hockey League (NHL) for a total of 2254 observations (i.e. teams in a given season). By using data aggregated over seasons we minimize the effect of proximal situational influences (home-field advantage, referee biases, etc.). As shown in Fig. 1, low ranking was consistently associated with foul ratio in each individual league as well as across leagues (rho=.30; Cohen's d=.63, p<.001). When pooling over all competitions (Fig. 1h), low-ranking teams (Z-transformed position>1)

showed on average a 9.5% greater foul ratio (104.68% vs 95.18%) than high-ranking ones (Z-transformed position<-1). A multiple linear regression revealed that the relationship between rank and foul ratio remained virtually unaltered (beta=.30, p<.001) when controlling for league and season, which had no effect and did not interact with ranking (all p>.345).

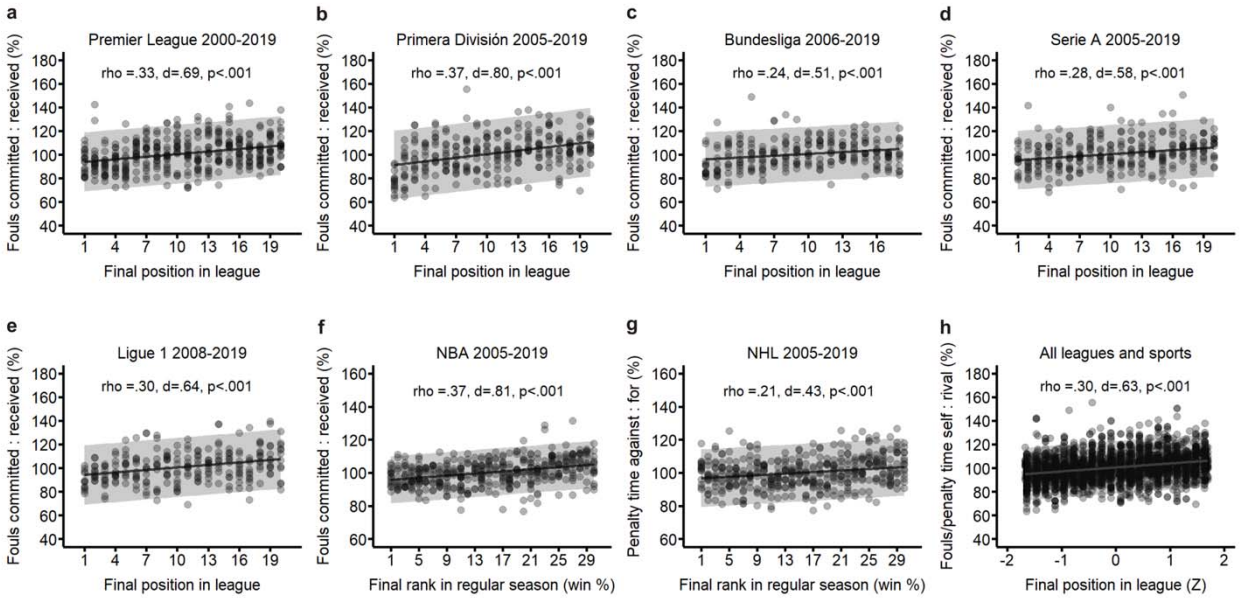


Fig. 1. Correlation between competitive status and aggression in elite sports teams. Status was defined as the team's final position at the end of a season. Aggression was defined as the ratio of fouls (soccer), personal fouls (basketball), or penalty time (ice hockey) incurred relative to received, in %. We collapsed data across seasons. **a**, English Premier League 2000-2019, **b**, Spanish Primera División 2005-2019, **c**, German Bundesliga 2006-2019, **d**, Italian Serie A 2005-2019, **e**, French Ligue 1 2008-2019, **f**, North-American National Basketball Association League 2005-2019, **g**, North-American National Basketball Association League 2005-2019, **h**, All data points collapsed, with position expressed as Z-scores.

Our analysis of sports data provides consistent -albeit purely correlational- evidence for a link between low competitive status and aggression at the level of teams. It is however unclear whether the status-aggression relationship can also be observed in individual competitors. Moreover, fouls in sports serve a predominantly instrumental function. A more stringent test of our main hypothesis would entail manipulating status experimentally during one-to-one contests in which aggression does not strictly fulfil a competitive purpose, i.e. it does not directly impact rivals' performance. To that end, we ran a pre-registered pilot study (https://osf.io/q5ge6) in order to probe whether individuals would be more or less aggressive against opponents varying in competitive status. Subjects were measured in groups of three and were told that they would compete against each other in a competitive reaction time task (Fig. 2a). Unbeknownst to them, the task was programmed so that they lost more often against one rival (66%,

high-status) than against the other (33%, low-status). On average, they won half of the trials and ended up in the middle rank. In each round of the task, the winner could punish the loser with a sound blast. Critically, participants set the volume of the sound blast at the beginning of each trial, which constitutes a reliable [27] and valid [28] measure of aggression. During punishment selection, the percentage of trials won by them and their opponent was presented as a filled bar on screen. Next came the reaction time task, in which they had to be quicker than their opponent in pressing a button when a jittered target appeared. In the outcome phase they were informed of whether they won or lost the contest and of the rival's selection. If they lost, they received the corresponding sound blast via headphones. Opponents' punishments were programmed to progressively increase over time in order to provoke participants.

Subjects were 25 men aged 24±3.1 (mean ± standard deviation) and reportedly free of current psychiatric and neurological conditions. Post-experimental manipulation checks revealed that they were aware of winning less often ($t=7.55$, $d=1.51$, $p<.001$) against the high-status player (34.8% ± 2.1% [mean ± standard error]) than against the low-status one (60.5% ± 1.9%). As intended, they did not perceive the high-status opponent to be more aggressive ($t=1.58$, $p=.125$). Following the pre-registered analysis plan (https://osf.io/q5ge6), we extracted mean punishment selections per condition and compared them with a paired t-test. Subjects were more aggressive ($t=2.66$, $d=.54$, $p=.013$; Fig. 2b) against the high- than against the low-status opponent.

In a subsequent neuroimaging study, we used a slightly modified version of the same task to scrutinize the neural response to status information during a competitive interaction. Further, we tested whether status-related brain activity was linked with actual aggressive behavior in the task as well as with salivary testosterone and cortisol measured before and after scanning. Unlike the pilot study, we programmed opponents' punishments to stay relatively constant over time so as to disentangle provocation and status effects. Participants were 47 healthy young men (age= 24 ± 3.8) who believed to be playing against two other participants. These were actually confederates and the task was again preprogramed. Here, the status manipulation was also successful, as subjects were aware of winning less often against the high- than against the low-status rival ($t=10.05$, $d=2.36$, $p<.001$). Subjects overestimated the percent of trials won against both the high- (41.1% ± 1.6% vs 33%, $t=4.74$, $d=.69$, $p<.001$) and low-status opponents (71.9% ± 1.7% vs 66%, $t=3.40$, $d=.49$, $p=.001$). Again, they did not perceive the high-status opponent to be more aggressive ($t=1.26$, $d=.27$, $p=.213$). As in the pilot study, subjects selected

higher average punishments against the high-status rival than against the low-status one (t=3.57, d=.57, p<.001; Fig. 2c). We thus calculated status-dependent aggression as the difference in average punishment selections against each opponent (high- minus low-status) to be used in later correlation analyses with neural and hormonal data.

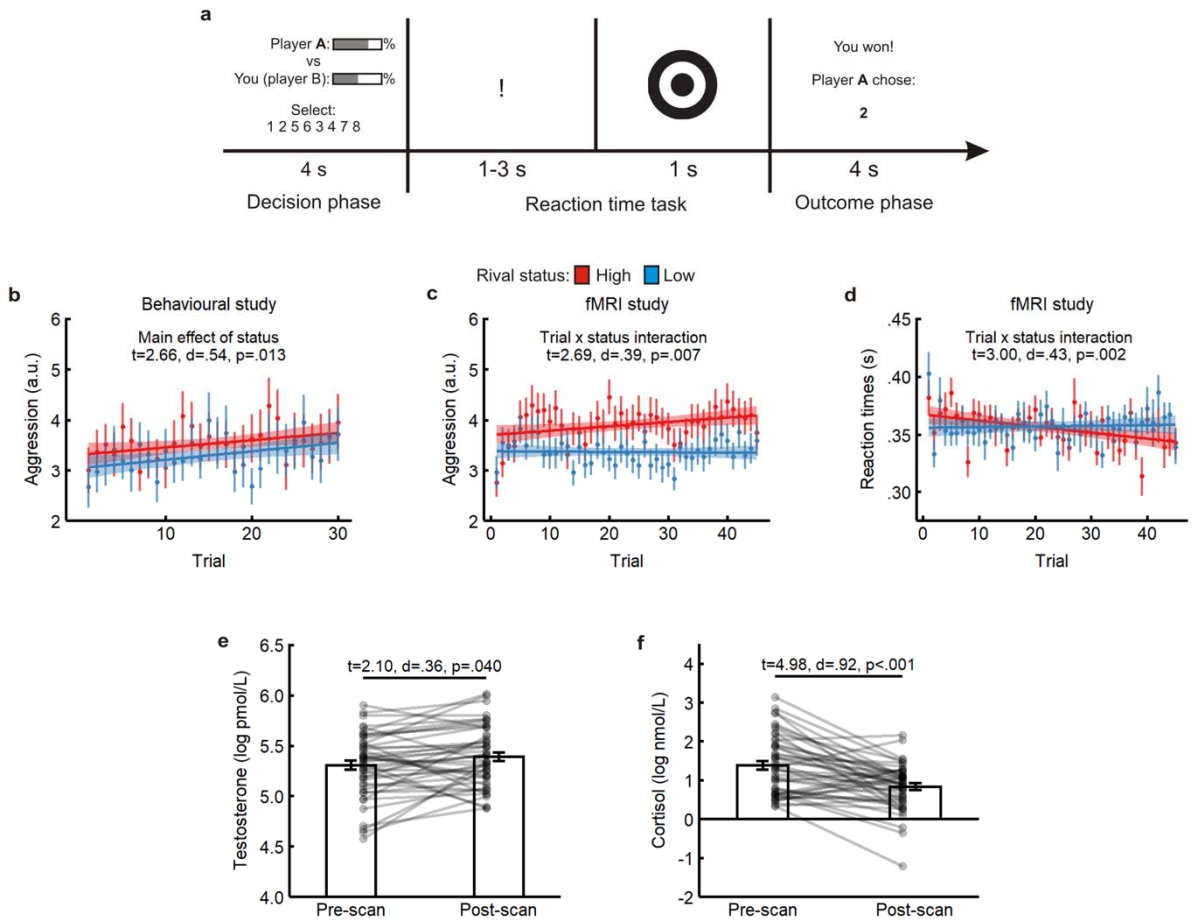


Fig. 2. a, outline of the task. Subjects saw the % of victories achieved by the opponent and themselves as a filled bar while choosing the volume of a sound blast to be later directed at their opponent (decision phase). Then, they had to be faster than the rival in pressing a button when a jittered bullseye appeared (reaction time task). In the outcome phase they were informed of whether they won or lost and of the rival's selection. If they lost, they received the sound blast at the end of the trial via headphones. One opponent was consistently better than them (high-status) and the other was worse (low-status), though both opponents chose equally strong sound blasts on average. **b**, mean aggression by trial and opponent in the pilot experiment, **c**, mean aggression by trial and opponent in the neuroimaging experiment. **d**, mean response latency in the reaction time task by trial and opponent in the neuroimaging experiment. **e**, testosterone levels in saliva before and after scanning. **f**, cortisol levels in saliva before and after scanning.

We additionally tested whether subjects' behavior changed over time using linear mixed-effects models (see Methods). Participants became linearly more aggressive against the high-status opponent,

whereas their aggression levels remain unchanged against the low-status one (status x trial interaction: t=2.69, d=.39, p=.007; Fig. 2c). Specifically, they evinced an increase in aggression of around .7 points (≈8.75% of the scale) in the last (4.09 ± .08) compared to the first five trials (3.39 ± .21) against the high-status opponent. In contrast, their average aggression levels remained virtually unchanged against the low-status rival (first five trials: 3.47 ± .14; last five trials: 3.49 ± .07). Thus, participants selected stronger punishments against the high-status opponent even when provocation remained constant.

Reaction times in the decision phase did not vary as a function of status, time, or their interaction (all p>.128). That is, subjects took similarly long to choose the punishment against either opponent across trials (high-status: 1321 ± 16ms; low-status: 1291 ± 12ms). Nevertheless, participants became quicker over time in the reaction time task against the high-status opponent (t=3.00, d=.43, p=.002; Fig. 2d). On average, subjects were approximately 22ms faster in the last five (350 ± 3ms) relative to the first five trials (372 ± 6ms) against the high-status opponent, whereas they were comparably quick throughout the task against the low-status rival (first five trials: 364 ± 12ms; last five trials: 365 ± 6ms). This confirms the motivational relevance of competitive status for participants' performance.

Concerning hormonal concentrations, testosterone increased after scanning (t=2.10, d=.36, p=.040; Fig. 2d) whereas cortisol decreased (t=4.96, d=.92, p<.001; Fig. 2e). Baseline concentrations of testosterone (r=.08, p=.595) or cortisol (r=-.01, p=.973) were not associated with aggression, nor was the testosterone-to-cortisol ratio or the pre-post change in either hormone (all p>.520). Therefore, we did not use endogenous hormone levels for further analyses.

With regards to neuroimaging data, we first tested whether regions-of-interest (ROIs) previously linked to the processing of status signals also encode competitive status during decisions to aggress (Fig. 3a). To that end, we performed multi-voxel pattern analysis (MVPA) in the decision phase (see Methods). This revealed that the striatum and hippocampus distinguished between opponents (Fig. 3b). That is, multivariate activity patterns in these regions could classify at an above-chance level whether subjects were facing the high- or the low-status opponent. Post-hoc searchlight analyses showed that both ventral and dorsal aspects of the left striatum contributed to the effect, whereas in the case of the hippocampus classification was most strongly driven by its right anterior portion (Fig. 3c). We then tested whether a stronger multivariate differentiation between opponents was linked with status-dependent aggression across subjects. To do so, we took classification performance values for the hippocampus and striatum

and correlated them with mean aggression against the high- minus low-status opponent. Although we observed no effects for the hippocampus (r=-.03, p=.823), there was an association between striatum activity patterns and aggression against the high- minus low-status rival (rho=.29, d=.62, p=.043; Fig. 3d). In other words, subjects showing a more pronounced differentiation between opponents in the striatum showed greater status-contingent aggressive behavior.

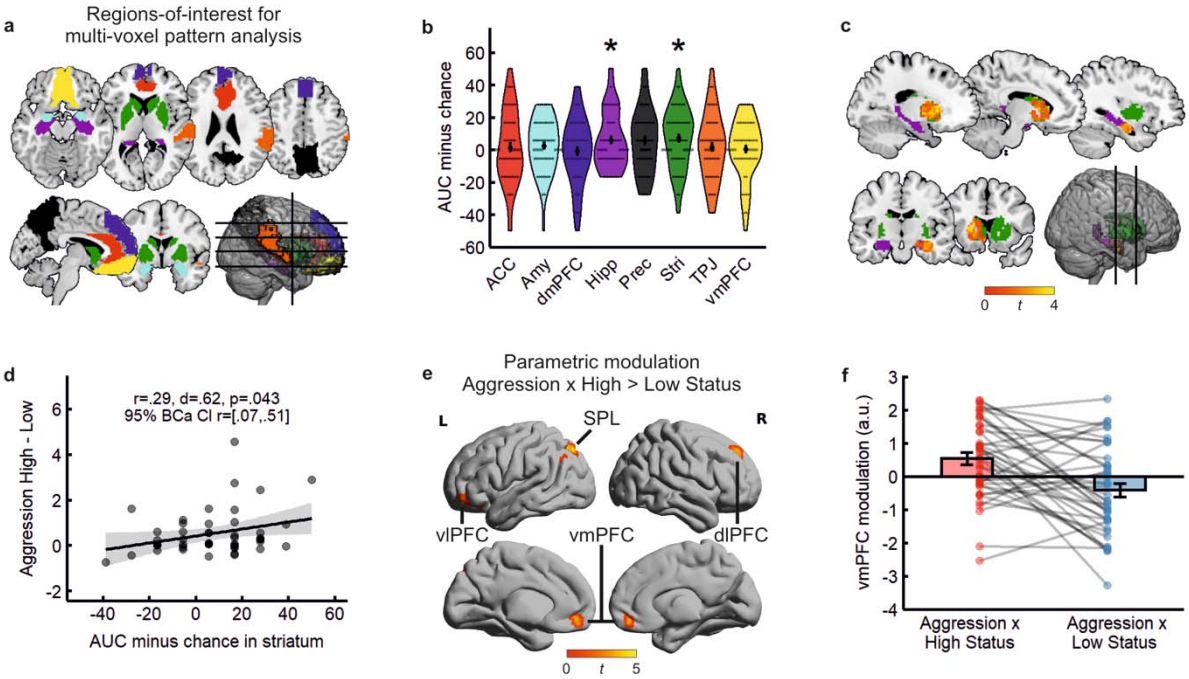


Fig. 3. Neural processing of competitive status during punishment selection (decision phase). **a**, regions-of-interest (ROIs) used in multi-voxel pattern analysis (MVPA) to identify multivariate status representations during punishment selection. **b**, classification performance expressed as area under the curve (AUC) minus chance per each ROI. Each dot shows the AUC value per subject, with higher values indicating better classification between high- and low-status rival. Asterisks denote ROIs with significant above-chance accuracy according to bootstrap resampling. ACC: anterior cingulate cortex; Amy: amygdala; dmPFC: dorsomedial prefrontal cortex; Hipp: hippocampus; Prec: precuneus; Stri: striatum; TPJ: temporal-parietal junction; vmPFC: ventromedial prefrontal cortex. **c**, results of the post-hoc searchlight MVPA analysis within hippocampus and striatum masks (voxel-level p<.05 uncorrected, cluster size k>50). Higher values indicate a greater contribution to the classification. **d**, correlation between classification performance in the striatum and mean aggression against the high- minus low-status opponent. r=Pearson correlation coefficient; d=Cohen's d; 95% BCa CI: 95% bias-corrected accelerated confidence intervals. **e**, parametric modulation in the decision phase (voxel-level p<.001, cluster-level p<.05 Family Wise-Error [FWE] corrected). Higher values indicate more trial-wise activity as a function of aggression against the high- minus low-status opponent. dlPFC: dorsolateral prefrontal cortex; SPL: superior parietal lobe; vlPFC: ventrolateral prefrontal cortex. **f**, parameter estimates resulting from the parametric modulation in the vmPFC. A.u.: arbitrary units.

We subsequently asked whether neural reactivity to status cues influenced decisions to aggress on a within-subject basis. We hence performed a parametric modulation analysis comparing trial-wise punishment selections against each opponent (see Methods). As shown in Fig. 3d, greater aggression

against the high- compared to the low-status opponent was associated with *increased* activity in ventromedial prefrontal cortex (vmPFC; t=4.57, k=59, x=9, y=50, z=-10; Fig. 3b), dorsolateral prefrontal cortex (dlPFC; t=4.11, k=57, x=21, y=44, z=41), ventrolateral prefrontal cortex (vlPFC; t=4.79, k=57, x=-33, y=32, z=-16) and superior parietal lobe (SPL; t=4.80, k=88, x=-24, y=-76, z=47). That is, activity in these regions co-varied with participants' aggressive behavior against the high- relative to the low-status opponent. No regions survived in the opposite contrast (low > high).

In the outcome phase, we observed widespread activation in the contrast won > lost with peaks in the bilateral ventral striatum (VS; Table S1 in Supplementary Material). This is in line with previous results employing competitive tasks [29,30]. More importantly, we tested the interaction between status and outcome in order to inspect whether victories and defeats were differently processed against the high- compared to the low-status opponent. The interaction contrast *[Won High + Lost Low)] > [Lost High + Won Low]* revealed activation in left anterior insula (AI), dorsal anterior cingulate cortex (dACC), left dorsolateral prefrontal cortex (dlPFC), left and right temporoparietal junction (TPJ), precuneus, and VS (Fig. 4a-c; see Table S1 in Supplementary Material for complete results). These regions hence reacted more strongly to infrequent, status-incongruent outcomes. However, as becomes apparent from the parameter estimates in the VS and dACC (Fig. 4b-c), the interaction was driven by a stronger effect of the outcome when playing against the high- relative to the low-status opponent. If the effect was merely driven by the infrequency or unpredictability of events, activity should be also increased in lost relative to won trials against the low-status opponent, which was not the case. The opposite contrast, which entailed expectable, status-congruent outcomes, yielded activation in primary visual and inferior temporal cortex as well as a cluster in the brainstem (see Table S1 in the Supplementary Material). We next probed whether the neural processing of victories and defeats was linked with status-dependent aggressive behavior. We therefore extracted parameter estimates from clusters resulting from the *[Won High + Lost Low)] > [Lost High + Won Low]* contrast (AI, dACC, dlPFC, left and right TPJ, precuneus, and VS) and correlated them with the status effect on aggression. Specifically, we extracted the high- minus low-status difference separately for victories and defeats. Because we performed 14 correlations (2 values for each of 7 ROIs), we applied a Bonferroni correction to control for multiple comparisons (.05/14=.003). Differential reactivity to victories in the dACC was significantly associated with aggression against the high- minus the low-status opponent (rho=.45, d=1.02, p=.001; Fig. 4d). There was no correlation for losses (r=.12, p=.394).

Activity in the rest of ROIs (AI, dlPFC, left and right TPJ, and VS) was not associated with status-dependent aggression at a corrected level (all p>.050).

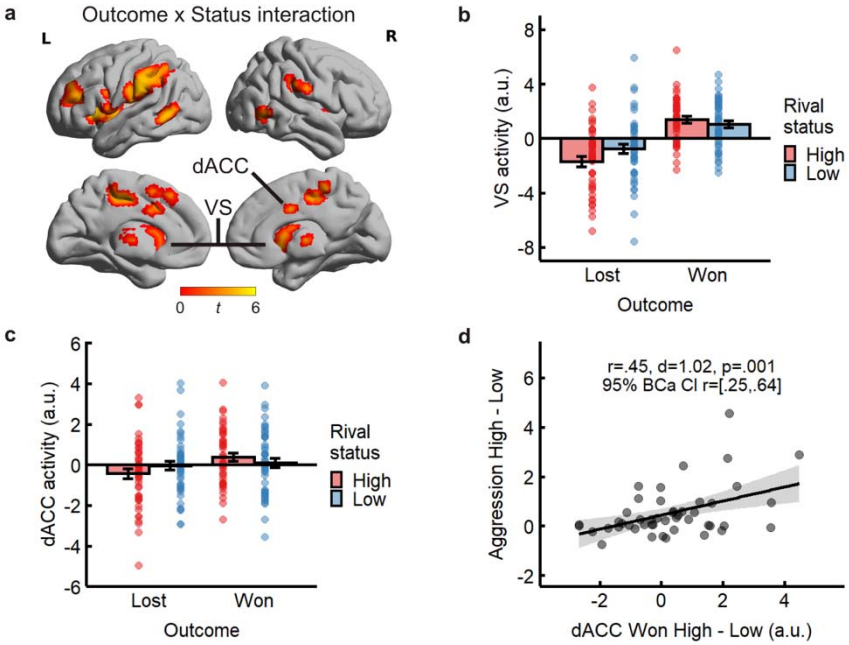


Fig. 4. Neural processing of competitive victories and defeats (outcome phase). **a**, regions showing a significant interaction between status and outcome (voxel-level p<.001, cluster-level p<.05 Family Wise-Error [FWE] corrected). **b**, parameter estimates for the interaction effect in the ventral striatum (VS). **c**, parameter estimates for the interaction effect in the dorsal anterior cingulate cortex (dACC). **d**, correlation between reactivity to victories and aggression against the high- minus low-status opponent in the dACC. r=Pearson correlation coefficient; d=Cohen's d; 95% BCa CI: 95% bootstrapped bias-corrected accelerated confidence intervals (BCa CI).

Discussion

Aggressiveness is often taken as a requisite for competitive success [1–6]. However, better contenders can rely on their skill, whereas less successful ones may use aggression to hamper rivals' progress [9,10]. Moreover, the frustration and loss of control entailed in being repeatedly defeated should facilitate rather than suppress aggression in low-status contestants [7,8,31]. In a correlational analysis and two laboratory experiments we found that low competitive status, measured as competitors' relative rank in a hierarchy, was linked with more frequent and intense aggressive behavior. Our results replicate and extend previous investigations showing that fouls or penalties are associated with worse overall team performance in American football [32], ice hockey [33], soccer [34], and basketball [35]. Unlike these studies, however, our measure of aggression (fouls or penalties committed relative to received) captures the reciprocal nature of aggressive play and thus provides clearer evidence on the status-aggression

relationship. The findings from our two experimental studies further show that the link between low status and aggression holds in competitions between individuals (i.e. not only between teams) and when aggressive behavior does not directly offer a clear competitive advantage. In addition, in our neuroimaging experiment subjects became quicker and more aggressive over time against a high-status rival, even when the opponents' hostile intent remained constant over time. This finding further demonstrates that aggression increases as competitive hierarchies emerge and is not a mere correlate of low performance. These results also imply that the relationship between competitive status and aggression is not limited to situations involving face-to-face competition, but also applies to virtual, indirect contests. In line with this assumption, an online bargaining study showed that individuals attribute more hostile intentions to higher-status opponents [36]. Competitive status thus appears to be a relevant trigger for aggressive behavior in both direct and virtual interactions.

Our main finding that participants behaved more aggressively when ranked in a low-status position is however not entirely expectable. After all, participants might have instead shown the opposite pattern, i.e. punishing a lower-status rival to assert their dominance while being less aggressive against a higher-ranking opponent as a conciliatory strategy. Indeed, some studies have shown that participants accept more unfair offers when they are in a low-status position [37,38]. Nevertheless, in these investigations behaving submissively yielded a direct economic benefit. What could explain the results of the present study, in which there was no money at stake? One possibility is that subjects were more willing to incur the risk of behaving aggressively when facing a high-status opponent, as the prospect of losing increases risk-seeking [39]. On the other hand, being in high-status position induces risk aversion, so that subjects might avoid punishing others due to e.g. reputational concerns [40]. Put otherwise, subjects had less to lose against the high-status opponent, which facilitates risk-taking. In addition, inducing a low-status mindset threatens self-worth and increases hostility [41], which might have further exacerbated status-dependent aggression in the present study.

We also observed a competition-induced increase in testosterone concentrations. This surge occurred instead of the decline that could be expected due to the hormone's circadian rhythm [42,43]. In men, testosterone tends to rise after winning [44] and when they achieve a higher status in a skill-based hierarchy [45]. Although subjects in the present study only won half of the trials on average, they did clearly defeat one opponent and overestimated the percentage of trials won against both rivals. Indeed,

elevations in testosterone levels have been linked to more positive self-appraisals of competitive performance [46]. These results therefore indicate that competition prompted changes in testosterone concentrations that are unaccounted for by naturally occurring, cyclic fluctuations. Cortisol concentrations on the other hand showed a steep decline after the task. This effect, which we have previously observed [42], is likely due to anticipatory arousal, circadian decline, and/or relaxation in the scanner; the latter might have been heightened by the post-task resting-state measurement. Nonetheless, contrary to some findings [21], neither endogenous testosterone nor cortisol were linked with status-dependent aggressive behavior. Our results are also in disagreement with the finding that competition-related surges in testosterone facilitate further antagonistic behavior [47,48]. Recent meta-analytic evidence suggests that the association between testosterone and aggression in humans is small [49]. Testosterone-by-cortisol interactions yield even lower effect sizes across studies, with highly variable estimates and some evidence for publication bias [50]. Large, pre-registered studies [51] are necessary to reach conclusive evidence on the role of testosterone in status-based aggression.

At the neural level, we identified a number of brain regions implicated in the processing of competitive hierarchies and characterized their relevance for status-dependent aggressive behavior. Multivariate analyses revealed that the right anterior hippocampus distinguished between the high- and low-status rival during punishment selection. A growing number of studies indicate that the role of the human anterior hippocampus is not limited to mapping events in time or items in space. Rather, this structure has been suggested to encode abstract relationships between entities along several dimensions extracted from statistical regularities in the environment, i.e. relational knowledge [52]. Similar mechanisms have been documented in the social domain, so that the anterior hippocampus tracks changes in an individual's power and affiliation relative to oneself [15,18]. Our results therefore provide converging evidence that this region, analogous to the ventral hippocampus in rodents [53], encodes status relationships and can thereby inform social decision-making. The left ventral and dorsal striatum also displayed multivariate signals distinguishing the high- and low-status rivals. Furthermore, participants with a stronger differentiation between opponents in the striatum showed more status-dependent aggression. Paralleling our findings, striatum neurons in the rhesus macaque brain respond differentially to social cues from dominant and submissive monkeys, which concurs with a preference to watch faces from higher-status individuals [54]. The present results thus indicate that the striatum does not only enable

the learning of skill-based hierarchies based on competitive outcomes [11], but also differentiates between individuals differing in status during decisions to aggress. Therefore, our data bridge animal and human research in showing that the hippocampus and striatum are involved in the relational processing of social dominance signals. The present findings additionally suggest that these mechanisms are at play during competitive interactions and might contribute to the adoption of aggressive strategies.

The ventromedial prefrontal cortex (vmPFC) evinced a status-dependent co-variation with participants' momentary changes in aggressive behavior. Put another way, this region integrated status information in decisions to aggress on a trial-by-trial level. A number of volumetric and functional studies indicate that the vmPFC is crucial for the control of aggressive impulses [55–57], presumably by downregulating amygdala activity [58]. There are however contradictory findings showing that vmPFC damage is linked with *increased* cooperative behavior [59], or that vmPFC activation *positively* predicts aggressive decisions [60,61]. Here, the status-contingent, aggression-related elevations in vmPFC activity went along with those of other regions putatively involved in emotion regulation and decision-making such as the dorsolateral or ventrolateral prefrontal cortex [62]. In contrast, there were no positive associations between brain structures assumed to generate aggressive urges (e.g. amygdala, periaqueductal grey) [31,58] and punishment selections in the task. Therefore, we speculate that the observed prefrontal and parietal activity might correspond to deliberate, intentional aggressive decisions. The relatively long and invariant response times during punishment selection (i.e. around 1300ms throughout the task) are in consonance with this interpretation. Future studies could explicitly address participants' strategies, as has been shown, for instance, in the domain of moral decision-making [63].

Finally, there was enhanced activation for status-incongruent outcomes in the VS, AI, dACC, dlPFC, TPJ and precuneus. More specifically, these regions showed a stronger differentiation between victories and defeats against the high- compared to the low-status rival. Our results resemble those from a previous study in which subjects competed against rivals varying in status [64], and further implicate the VS in competitive hierarchy learning [12]. The present findings are also consistent with the observation that brain areas involved in valuation (e.g. VS) and social cognition (e.g. TPJ, precuneus) display differential sensitivity to a person's position in a hierarchy [16]. In addition, dACC reactivity to victories against the high- minus the low-status rival predicted greater aggression against the former relative to the latter. That is, reactivity to status-enhancing wins, as compared to status-asserting ones, was correlated

with greater status-dependent aggression. Although the precise role of the dACC remains a contentious issue, a recent investigation found that surprise detection best explains dACC activity during decision-making as compared to other formal accounts [65]. In the context of social cognition, similar sequential updating mechanisms in the cingulate gyrus have been postulated to track others' motivation during social interactions [66] as well as their rank and power within a hierarchy [15]. Taken together, these results tentatively suggest that status-based aggression might be particularly pronounced in subjects for whom status-enhancing victories are more salient.

In sum, we found that low competitive status consistently evokes aggression in healthy young men. Our findings refute the widespread notion that successful competitors are the more aggressive ones. Rather, our results indicate that those in the lower echelons of competitive hierarchies use aggression more frequently and intensely over the course of successive competitions. Our experimental data shows that this holds even for situations in which aggression does not directly serve any instrumental function. Although endogenous testosterone levels increased after a competitive interaction, there were no links between this hormone and aggression, neither alone nor in interaction with cortisol. The latter finding questions a strong impact of testosterone on status-driven aggression, in line with recent meta-analytic evidence [49,50]. On a neural level, both anterior hippocampus and striatum displayed multivariate representations of competitive status. Interindividual differences in status-dependent aggressive behavior could be predicted by a stronger differentiation between opponents in the striatum and by neural reactivity to status-enhancing victories in the dACC. On a trial-by-trial, within-subject basis, the vmPFC together with other prefrontal and parietal regions favored status-based aggression. In sum, our study reveals a consistent association between low competitive status and aggression, and points toward potential neural mechanisms underlying this relationship.

References

1. Chow, G.M., Murray, K.E., and Feltz, D.L. (2009). Individual, Team, and Coach Predictors of Players' Likelihood to Aggress in Youth Soccer. *31*, 425.
2. Trebicky, V., Havlícek, J., Roberts, S.C., Little, A.C., and Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. Psychol. Sci. *24*, 1664–1672.
3. Weick, M. (2020). Power and aggression: making sense of a fickle relationship. Curr. Opin. Psychol. *33*, 245–249.
4. Cheng, J.T. (2020). Dominance, prestige, and the role of leveling in human social hierarchy and equality. Curr. Opin. Psychol. *33*, 238–244.

5. Waasdorp, T.E., Baker, C.N., Paskewich, B.S., and Leff, S.S. (2013). The association between forms of aggression, leadership, and social status among urban youth. J. Youth Adolesc. *42*, 263–274.
6. Laustsen, L., and Petersen, M.B. (2017). Perceived Conflict and Leader Dominance: Individual and Contextual Factors Behind Preferences for Dominant Leaders. Polit. Psychol. *38*, 1083–1101.
7. Przybylski, A.K., Deci, E.L., Rigby, C.S., and Ryan, R.M. (2014). Competence-impeding electronic games and players' aggressive feelings, thoughts, and behaviors. J. Pers. Soc. Psychol. *106*, 441–457.
8. Oxford, J.K., Tiedtke, J.M., Ossmann, A., Özbe, D., and Schultheiss, O.C. (2017). Endocrine and aggressive responses to competition are moderated by contest outcome, gender, individual versus team competition, and implicit motives. PLoS ONE *12*, e0181610.
9. Coulomb-Cabagno, G., and Rascle, O. (2006). Team Sports Players' Observed Aggresion as a Function of Gender, Competitive Level, and Sport Type. J. Appl. Soc. Psychol. *36*, 1980–2000.
10. Kirker, B., Tenenbaum, G., and Mattson, J. (2000). An Investigation of the Dynamics of Aggression: Direct Observations in Ice Hockey and Basketball. Res. Q. Exerc. Sport *71*, 373–386.
11. Qu, C., Ligneul, R., Van der Henst, J.-B., and Dreher, J.-C. (2017). An Integrative Interdisciplinary Perspective on Social Dominance Hierarchies. Trends Cogn. Sci. *21*, 893–908.
12. Ligneul, R., Obeso, I., Ruff, C.C., and Dreher, J.-C. (2016). Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex. Curr. Biol. *26*, 3107–3115.
13. Hassabis, D., Spreng, R.N., Rusu, A.A., Robbins, C.A., Mar, R.A., and Schacter, D.L. (2013). Imagine All the People: How the Brain Creates and Uses Personality Models to Predict Behavior. Cereb. Cortex *24*, 1979–1987.
14. Kumaran, D., Melo, H.L., and Duzel, E. (2012). The emergence and representation of knowledge about social and nonsocial hierarchies. Neuron *76*, 653–666.
15. Kumaran, D., Banino, A., Blundell, C., Hassabis, D., and Dayan, P. (2016). Computations Underlying Social Hierarchy Learning: Distinct Neural Mechanisms for Updating and Representing Self-Relevant Information. Neuron *92*, 1135–1147.
16. Zerubavel, N., Bearman, P.S., Weber, J., and Ochsner, K.N. (2015). Neural mechanisms tracking popularity in real-world social networks. Proc. Natl. Acad. Sci. *112*, 15072.
17. Schafer, M., and Schiller, D. (2018). Navigating Social Space. Neuron *100*, 476–489.
18. Tavares, R.M., Mendelsohn, A., Grossman, Y., Williams, C.H., Shapiro, M., Trope, Y., and Schiller, D. (2015). A Map for Social Navigation in the Human Brain. Neuron *87*, 231–243.
19. Muscatell, K.A., Morelli, S.A., Falk, E.B., Way, B.M., Pfeifer, J.H., Galinsky, A.D., Lieberman, M.D., Dapretto, M., and Eisenberger, N.I. (2012). Social status modulates neural activity in the mentalizing network. NeuroImage *60*, 1771–1777.
20. Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., and Greicius, M.D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. J. Neurosci. Off. J. Soc. Neurosci. *27*, 2349–2356.
21. Carré, J.M., and Olmstead, N.A. (2015). Social neuroendocrinology of human aggression: Examining the role of competition-induced testosterone dynamics. Neuroscience *286*, 171–186.
22. Terburg, D., and van Honk, J. (2013). Approach–Avoidance versus Dominance–Submissiveness: A Multilevel Neural Framework on How Testosterone Promotes Social Status. Emot. Rev. *5*, 296–302.
23. Inoue, Y., Takahashi, T., Burriss, R.P., Arai, S., Hasegawa, T., Yamagishi, T., and Kiyonari, T. (2017). Testosterone promotes either dominance or submissiveness in the Ultimatum Game depending on players' social rank. Sci. Rep. *7*, 5335.
24. Dreher, J.-C., Dunne, S., Pazderska, A., Frodl, T., Nolan, J.J., and O'Doherty, J.P. (2016). Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. Proc. Natl. Acad. Sci. *113*, 11633–11638.
25. Josephs, R.A., Sellers, J.G., Newman, M.L., and Mehta, P.H. (2006). The Mismatch Effect: When Testosterone and Status Are at Odds. J. Pers. Soc. Psychol. *90*, 999–1013.
26. Casto, K.V., and Edwards, D.A. (2016). Testosterone, cortisol, and human competition. Horm. Behav. *82*, 21–37.
27. Chester, D.S., and Lasko, E.N. (2019). Validating a Standardized Approach to the Taylor Aggression Paradigm. Soc. Psychol. Personal. Sci. *10*, 620–631.
28. King, A.R., and Russell, T.D. (2019). Lifetime Acts of Violence Assessment (LAVA) predictors of laboratory aggression. Aggress. Behav. *45*, 477–488.

29. Votinov, M., Pripfl, J., Windischberger, C., Sailer, U., and Lamm, C. (2015). Better you lose than I do: neural networks involved in winning and losing in a real time strictly competitive game. Sci. Rep. *5*, 11017.
30. Buades-Rotger, M., Brunnlieb, C., Münte, T.F., Heldmann, M., and Krämer, U.M. (2016). Winning is not enough: ventral striatum connectivity during physical aggression. Brain Imaging Behav. *10*, 105–114.
31. Yu, R., Mobbs, D., Seymour, B., Rowe, J.B., and Calder, A.J. (2014). The neural signature of escalating frustration in humans. Cortex *54*, 165–178.
32. Hauge, J.A. (2012). Incentive for Aggression in American Football. In Violence and Aggression in Sporting Contests: Economics, History and Policy, R. T. Jewell, ed. (New York, NY: Springer New York), pp. 29–46. Available at: https://doi.org/10.1007/978-1-4419-6630-8_3.
33. Coates, D., Battré, M., and Deutscher, C. (2012). Does Violence in Professional Ice Hockey Pay? Cross Country Evidence from Three Leagues. In Violence and Aggression in Sporting Contests: Economics, History and Policy, R. T. Jewell, ed. (New York, NY: Springer New York), pp. 47–63. Available at: https://doi.org/10.1007/978-1-4419-6630-8_4.
34. Jewell, R.T. (2012). Aggressive Play and Demand for English Premier League Football. In Violence and Aggression in Sporting Contests: Economics, History and Policy, R. T. Jewell, ed. (New York, NY: Springer New York), pp. 113–131. Available at: https://doi.org/10.1007/978-1-4419-6630-8_8.
35. Berri, D.J., and Rodenberg, R.M. (2012). Crime and Punishment in the National Basketball Association. In Violence and Aggression in Sporting Contests: Economics, History and Policy, R. T. Jewell, ed. (New York, NY: Springer New York), pp. 65–76. Available at: https://doi.org/10.1007/978-1-4419-6630-8_5.
36. Saalfeld, V., Ramadan, Z., Bell, V., and Raihani, N.J. (2018). Experimentally induced social threat increases paranoid thinking. R. Soc. Open Sci. *5*. Available at: http://rsos.royalsocietypublishing.org/content/royopensci/5/8/180569.full.pdf.
37. Blue, P.R., Hu, J., Wang, X., van Dijk, E., and Zhou, X. (2016). When Do Low Status Individuals Accept Less? The Interaction between Self- and Other-Status during Resource Distribution. Front. Psychol. *7*. Available at: https://www.frontiersin.org/article/10.3389/fpsyg.2016.01667.
38. Hu, J., Blue, P.R., Yu, H., Gong, X., Xiang, Y., Jiang, C., and Zhou, X. (2016). Social status modulates the neural response to unfairness. Soc. Cogn. Affect. Neurosci. *11*, 1–10.
39. Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science *211*, 453–458.
40. Dreber, A., Rand, D.G., Fudenberg, D., and Nowak, M.A. (2008). Winners don't punish. Nature *452*, 348–351.
41. Davis, J.R., and Reyna, C. (2015). Seeing red: How perceptions of social status and worth influence hostile attributions and endorsement of aggression. Br. J. Soc. Psychol. *54*, 728–747.
42. Buades-Rotger, M., Engelke, C., Beyer, F., Keevil, B.G., Brabant, G., and Krämer, U.M. (2016). Endogenous testosterone is associated with lower amygdala reactivity to angry faces and reduced aggressive behavior in healthy young women. Sci. Rep. *6*, 38538.
43. Keevil, B., MacDonald, P., Macdowall, W., Lee, D., and Wu, F. (2013). Salivary testosterone measurement by liquid chromatography tandem mass spectrometry in adult males and females. Ann. Clin. Biochem. Int. J. Biochem. Lab. Med. *51*, 368–378.
44. Geniole, S.N., Bird, B.M., Ruddick, E.L., and Carré, J.M. (2017). Effects of competition outcome on testosterone concentrations in humans: An updated meta-analysis. Horm. Behav. *92*, 37–50.
45. Cheng, J.T., Kornienko, O., and Granger, D.A. (2018). Prestige in a large-scale social group predicts longitudinal changes in testosterone. J. Pers. Soc. Psychol. *114*, 924–944.
46. Casto, K.V., Rivell, A., and Edwards, D.A. (2017). Competition-related testosterone, cortisol, and perceived personal success in recreational women athletes. Horm. Behav. Available at: http://www.sciencedirect.com/science/article/pii/S0018506X1730185X.
47. Carré, J.M., Campbell, J.A., Lozoya, E., Goetz, S.M.M., and Welker, K.M. (2013). Changes in testosterone mediate the effect of winning on subsequent aggressive behaviour. Psychoneuroendocrinology *38*, 2034–2041.
48. Zilioli, S., and Watson, N.V. (2014). Testosterone across successive competitions: Evidence for a 'winner effect' in humans? Psychoneuroendocrinology *47*, 1–9.
49. Geniole, S.N., Bird, B.M., McVittie, J.S., Purcell, R.B., Archer, J., and Carré, J.M. (2019). Is testosterone linked to human aggression? A meta-analytic examination of the relationship between baseline, dynamic, and manipulated testosterone on human aggression. Horm. Behav., 104644.

50. Dekkers, T.J., van Rentergem, J.A.A., Meijer, B., Popma, A., Wagemaker, E., and Huizenga, H.M. (2019). A meta-analytical evaluation of the dual-hormone hypothesis: Does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? Neurosci. Biobehav. Rev. *96*, 250–271.
51. Brannon, S.M., Carr, S., Jin, E.S., Josephs, R.A., and Gawronski, B. (2019). Exogenous testosterone increases sensitivity to moral norms in moral dilemma judgements. Nat. Hum. Behav. *3*, 856–866.
52. Garvert, M.M., Dolan, R.J., and Behrens, T.E.J. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. eLife *6*, e17086.
53. Fanselow, M.S., and Dong, H.-W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? Neuron *65*, 7–19.
54. Klein, J.T., and Platt, M.L. (2013). Social information signaling by neurons in primate striatum. Curr. Biol. CB *23*, 691–696.
55. Chester, D.S., Lynam, D.R., Milich, R., and DeWall, C.N. (2017). Physical aggressiveness and gray matter deficits in ventromedial prefrontal cortex. Cortex *97*, 17–22.
56. Gilam, G., Lin, T., Raz, G., Azrielant, S., Fruchter, E., Ariely, D., and Hendler, T. (2015). Neural substrates underlying the tendency to accept anger-infused ultimatum offers during dynamic social interactions. NeuroImage *120*, 400–411.
57. Beyer, F., Münte, T.F., Göttlich, M., and Krämer, U.M. (2015). Orbitofrontal Cortex Reactivity to Angry Facial Expression in a Social Interaction Correlates with Aggressive Behavior. Cereb. Cortex *25*, 3057–3063.
58. Coccaro, E.F., Sripada, C.S., Yanowitch, R.N., and Phan, K.L. (2011). Corticolimbic Function in Impulsive Aggressive Behavior. Biol. Psychiatry *69*, 1153–1159.
59. Wills, J., FeldmanHall, O., NYU PROSPEC Collaboration, Meager, M.R., and Van Bavel, J.J. (2018). Dissociable contributions of the prefrontal cortex in group-based cooperation. Soc. Cogn. Affect. Neurosci. *13*, 349–356.
60. Repple, J., Pawliczek, C.M., Voss, B., Siegel, S., Schneider, F., Kohn, N., and Habel, U. (2017). From provocation to aggression: the neural network. BMC Neurosci. *18*, 73.
61. Buades-Rotger, M., Beyer, F., and Krämer, U.M. (2017). Avoidant Responses to Interpersonal Provocation Are Associated with Increased Amygdala and Decreased Mentalizing Network Activity. eneuro *4*. Available at: https://doi.org/10.1523/ENEURO.0337-16.2017.
62. Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., and Heekeren, H.R. (2015). Changes in Effective Connectivity Between Dorsal and Ventral Prefrontal Regions Moderate Emotion Regulation. Cereb. Cortex *26*, 1923–1937.
63. van Baar, J.M., Chang, L.J., and Sanfey, A.G. (2019). The computational and neural substrates of moral strategies in social decision-making. Nat. Commun. *10*, 1483.
64. Zink, C.F., Tong, Y., Chen, Q., Bassett, D.S., Stein, J.L., and Meyer-Lindenberg, A. (2008). Know Your Place: Neural Processing of Social Hierarchy in Humans. Neuron *58*, 273–283.
65. Vassena, E., Deraeve, J., and Alexander, W.H. (2020). Surprise, value and control in anterior cingulate cortex during speeded decision-making. Nat. Hum. Behav. Available at: https://doi.org/10.1038/s41562-019-0801-5.
66. Apps, M.A.J., Rushworth, M.F.S., and Chang, S.W.C. (2016). The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. Neuron *90*, 692–707.
67. Perogamvros, I., Owen, L.J., Newell-Price, J., Ray, D.W., Trainer, P.J., and Keevil, B.G. (2009). Simultaneous measurement of cortisol and cortisone in human saliva using liquid chromatography–tandem mass spectrometry: Application in basal and stimulated conditions. J. Chromatogr. B *877*, 3771–3775.
68. Hebart, M.N., Görgen, K., and Haynes, J.-D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. Front. Neuroinformatics *8*. Available at: https://www.frontiersin.org/article/10.3389/fninf.2014.00088.
69. Maldjian, J.A., Laurienti, P.J., Kraft, R.A., and Burdette, J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. NeuroImage *19*, 1233–1239.
70. Nichols, T.E., Das, S., Eickhoff, S.B., Evans, A.C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M.P., Poldrack, R.A., Poline, J.-B., *et al.* (2017). Best practices in data analysis and sharing in neuroimaging using MRI. Nat. Neurosci. *20*, 299–303.

Methods

Correlational study

We extracted the ratio of fouls (soccer), personal fouls (basketball), or penalties (ice hockey) committed relative to those received for each team across the whole season, as well as each team's position at the end of the regular season. For basketball and ice hockey, we determined the position of teams from different divisions on the basis of their win percentage. We manually extracted data from http://www.footstats.co.uk (soccer), http://www.nba.com (basketball), and http://www.hockey-reference.com (ice hockey) since the last year in which complete, correct, data was available onwards (2000-2008 for European football leagues, 2005 for basketball and hockey). The correlation between received and committed fouls or penalties was r=.74 for football, r=.70 for basketball and r=.93 for ice hockey. This indicates that a) aggressive play follows reciprocal, "give-and-take" dynamics so that teams who commit more fouls also tend to receive more, but b) that the ratio of fouls/penalties committed to received captures some unique variance in teams' aggressiveness that is not accounted for by the raw amount of fouls or penalty time incurred. This is thus a valid measure to assess the degree to which a team engages in sanctioned aggression. Because of the ranked nature of the data, we computed Spearman correlation coefficients between position and foul ratio. We also tested whether this relationship varied across seasons and leagues with a multiple linear regression. To do so, we defined season and league as numerical predictors and defined the main effects and interactions between these two variables as well as rank (Z-transformed). All analyses were performed in R version 3.6.1 running on R Studio 1.1.423. We used the psych package for statistical analysis and the ggplot2 package for plotting. Data and analysis scripts for this analysis are available via the Open Science Foundation (https://osf.io/2jvx4/).

Behavioral experiment

In order to investigate the effect of status on aggression experimentally, we devised a modified version of the Taylor Aggression Paradigm, a competitive reaction time task in which the winner can choose the volume of a sound blast to punish the loser [42]. As this was an exploratory study, we aimed for a sample size of n=30 in order to have 80% power to detect a middle-sized within-subject effect (d=.5) at a conventional two-sided threshold of p<.05. After exclusion of three participants (two saw through the status manipulation and one failed to understand the task), the sample comprised 25 healthy young men (age: 24±3.1; height: 180±5.1 cm; weight: 75.96±9 kg; 23 right-handed, 2 ambidextrous). Participants

were measured in groups of three and believed they would compete against each other, but in reality they all played against the computer. A male confederate filled in for a participant when an appointment with three subjects was not possible. We first took participants to a computer room in which they provided informed consent and read the instructions together. Computers were separated by screen-walls so that participants could not see each other while playing, and they wore headphones throughout. We programmed the task so that participants competed in random order, but not more than three times in a row, against each opponent. After eight practice trials, the task began (see Fig. 2a for an outline of the task). Trials began with a decision phase wherein subjects saw the percentage of trials won by their current opponent and by themselves as separate filled bars, i.e. their competitive status. During the decision phase, participants set the volume of a sound blast in a 1-8 scale to be later delivered at their opponent should they win a subsequent reaction time task. The chosen volume was our measure of aggression. In the reaction time task, they had to be quicker than their rival in pressing any button when a target appeared (i.e. a bullseye). We manipulated status so that participants lost more frequently against one rival (66%, high-status) than against the other (33%, low-status) for a total of 60 trials. To make the setting more believable, subjects were more likely to lose when they were slower than their own median cumulative reaction time, and we interspersed shorter (4s) and longer (12s) inter-trial intervals; during the latter, the opponents allegedly played against each other. Following previous work [42], we programmed the opponents' punishment selections to increase progressively over time (i.e. every 20 trials) in order to gradually provoke participants.

Average punishment selections show high internal consistency, load onto a single factor [27], and predict real-life aggression [28], suggesting that they are a valid measure of aggression and that they adequately summarize participant's behavior in the task. We followed the pre-registered analysis plan (https://osf.io/q5ge6) and compared mean aggression against the high- vs low-status opponent with a paired t-test in R (version 3.6.1) running on R Studio (version 1.1.423). The study had been approved by the Ethics Committee of the University of Lübeck. Subjects consented to participate and to the anonymized dissemination of the data. They were debriefed regarding the goals and methods of the study. Data and analysis scripts for this study are freely available via the Open Science Foundation (https://osf.io/2jvx4/).

Neuroimaging experiment

In the follow-up neuroimaging study, we inspected how competitive status cues are processed in the context of an aggressive interaction. Though the task was almost identical to the pilot experiment, we here programmed the opponents to select punishments in the middle range (3-6) throughout the task instead of increasing progressively. We did so in order to disambiguate the effects of provocation and status, and to render trials more comparable throughout the task for subsequent averaging of brain activity. We also lengthened the task for a total of 90 trials, divided in 3 runs of 30 trials each (15 per opponent), so as to increase statistical power. In this case, the opponents were always confederates of the experimenters who pretended to play against the participant from computers outside the scanner. To increase the plausibility of this setting, participants played the eight practice trials together with the confederates in the same computer room where the pilot measurements had taken place.

We determined the sample size on the basis of a previous study (n=39) in which we observed middle-sized correlations (r≈.4) between endogenous testosterone, aggression, and brain activity using a similar task [42]. We thus aimed for a sample size of 50, which yields ~80% power to detect effects of r=.4 or larger with a two-sided threshold of p<.05. After exclusion of six subjects due to non-deception (n=4) and failure to understand the task (n=2), the final sample comprised 47 men aged 24±3.8 (height: 182±6.6 cm; weight: 79.90±11 kg; 37 right-handed, 7 left-handed, 3 ambidextrous) who were reportedly free of current psychiatric, neurological, and endocrine diseases except for one participant with treated hypothyroidism. Measurements began between 12 and 15 PM in order to minimize circadian fluctuations in hormonal levels, with the exception of two measurements (17 and 18 PM) that could not be scheduled otherwise.

Subjects provided saliva samples in plastic vials (SafeSeal micro tube 2ml from Sarstedt) with the passive drooling technique before and after scanning. We froze the samples at -20°C and shipped them in dry ice to author BGK's laboratory in Manchester (UK) for analysis once study was completed. Free testosterone and cortisol concentrations were estimated with liquid chromatography tandem mass spectrometry (LC-MS/MS) as described elsewhere [43,67]. Coefficients of variation (CV) with this technique have been reported to be 5.3% for testosterone and 8.7% for cortisol, whereas mean inter-assay CV were 9% for testosterone and 7.8% for cortisol [43,67]. The lower limits of quantification (LLOQ) were 5 pmol/L for testosterone and 0.8 nmol/L for cortisol.

We acquired all scans with a 64-channel head-coil mounted on a 3-T Siemens Magnetom Skyra scanner at the Center for Brain, Behavior and Metabolism (CBBM) at the University of Lübeck. For functional volumes we applied a single-shot gradient-recalled echo-planar imaging (GRE-EPI) sequence sensitive to blood oxygen level dependent (BOLD) contrast (TR=1060 ms; TE=30 ms; flip angle=60°; 60 transversal slices; slice thickness 3 mm; in-plane voxel size=3×3 mm^2; field-of-view [FOV]=210×210 mm^2, simultaneous multi-slice factor=4; full-brain coverage). We acquired 570 volumes per run for a total of three runs (≈30 minutes). Before and after the task we acquired two magnetic field measurements (2D double-echo gradient-echo sequence; TE_1=5,17ms; TE_2=7.63ms; TR=554ms; flip angle=60°; 50 transversal slices; slice thickness 3 mm; in-plane voxel size=3×3 mm^2; FOV=240x240 mm^2) for subsequent offline distortion correction as well as two 6.5 minutes eyes-closed resting-state runs not analysed for the present manuscript. We further acquired an anatomical scan at the start of each measurement applying a 3D T1-weighted Magnetization Prepared-RApid Gradient Echo (MP-RAGE) sequence (TR=2300 ms; TE=2.94 ms; TI=900 ms; flip angle=9°; voxel size=1×1×1 mm^3; FOV=320x320mm^2).

For the analysis of behavioral data, we first ran a paired t-test comparing mean aggression against the high- vs low-status opponent as we had done in the pilot experiment. Additionally, we ran linear mixed-effects models on participants' aggressive behavior and reaction times in the decision and outcome phases. We specifically tested whether aggression varied as a function of opponent status (high or low) and trial number (1 to 45). These two variables were defined as fixed effects whereas subject was defined as random effect. We performed these analyses using the lmerTest package in R (version 3.6.1) running on R Studio (version 1.1.423). For testosterone and cortisol, we compared the pre- and post-scan log-transformed values using paired t-tests in order to inspect for state changes in hormonal concentrations. Two pre-scan testosterone samples had to be discarded because the amount of saliva did not suffice for reliable quantification. We then tested whether mean aggression against the high- minus the low-status opponent was associated with baseline testosterone, cortisol, the testosterone-to-cortisol ratio, and/or pre-post changes in either hormone using Pearson correlation coefficients.

We performed the pre-processing and analysis of neuroimaging data using Statistical Parametric Mapping 12 (SPM12; https://www.fil.ion.ucl.ac.uk/spm/) running on Matlab 2019b. We used a standard pre-processing pipeline that involved, in this order, slice-timing correction to the four middle slices (i.e.

those acquired at 485ms) with a 4th degree spline interpolation, realignment to the first functional volume, coregistration of mean functional and anatomical images, segmentation of the anatomical images based on default tissue probability maps, normalization to the Montreal Neurological Institute (MNI) template and smoothing with an 8mm full-width at half-maximum Gaussian kernel. Additionally, we performed distortion correction using the FieldMap toolbox, except in seven participants for which this procedure failed or worsened image quality. We excluded one imaging run in three subjects due to excessive head movements (>3mm or 3° in any direction).

First-level models comprised two regressors for the decision phase (high- or low-status opponent; 4s) and four for the outcome phase (won or lost vs high- or low-status; 4s). We also included the absolute distance in status between subject and opponent in percentage as a covariate in the decision phase to control for momentary fluctuations in status. As regressors of no interest, we modelled the warning sign, target, and motor responses in the reaction time task as well as the sound at the end of the outcome phase as delta functions with null duration. We further included the six movement parameters estimated from realignment (x, y, z, pitch, roll, and yaw). We convolved all regressors (except for motion parameters) with the canonical hemodynamic response function, implemented a high-pass filter with a 128s cut-off, and applied the standard SPM autoregressive model AR(1) to control for autocorrelation. In order to inspect for trial-wise changes in activity associated with aggression, we additionally defined identical first-level models that included trial-wise mean-centered punishment selections as parametric modulator.

We first performed multi-voxel pattern analyses (MVPA) to uncover local activation patterns that most differentiated between the two opponents in the decision phase. We used the Decoding Toolbox [68] running on Matlab 2019b to implement a linear support vector machine algorithm that classifies multivariate activation patterns for each condition and run using a leave-one-out cross validation procedure on the corresponding beta images. We ran such an analysis within each of eight regions-of-interest (ROIs) putatively involved in the relational processing of social status [11,16,17]. We defined these ROIs as bilateral anatomical masks extracted from the Automatic Anatomical Labelling atlas [69]: the amygdala (homonymous mask), anterior cingulate cortex (“Cingulum_Ant” mask), dorsomedial prefrontal cortex (“Frontal_Sup_Medial” mask), hippocampus (homonymous mask), precuneus (homonymous mask), striatum (merge of “Caudate”, “Putamen”, and “Pallidum” masks), temporal-parietal junction (merge of “Temporal_Sup” and “SupraMarginal” masks), and ventromedial prefrontal cortex

(merge of “Rectus” and “Frontal_Med_Orb” masks). We resampled the masks to match the voxel size of the functional data (3mm). For each of these regions, we extracted the participant-wise area under the curve (AUC) minus chance as a classification performance measure. Here, AUC expresses the total surface below the receiver operating characteristic curve, which is obtained by plotting the cumulative true positive rates against the cumulative false positive rates. Hence, higher AUC values indicate a greater capacity to classify between categories. While standard accuracy measures treat all classifications equally, AUC also takes into account the “confidence” (i.e. the distance to the decision boundary) that the classifier has regarding the membership of specific items [68]. In order to assess statistically which ROIs distinguished between conditions, we computed the bootstrapped 95% bias-corrected accelerated confidence intervals (BCa CI) confidence intervals of the average AUC for each ROI. ROIs whose CI did not include zero were considered to classify between the high- and low-status rival at an above-chance level. In order to more precisely localize the observed effects, we performed a post-hoc searchlight analysis (9-voxel radius sphere) within each ROI showing above-chance classification. These analyses were thresholded at $p<.05$, $k>50$ uncorrected.

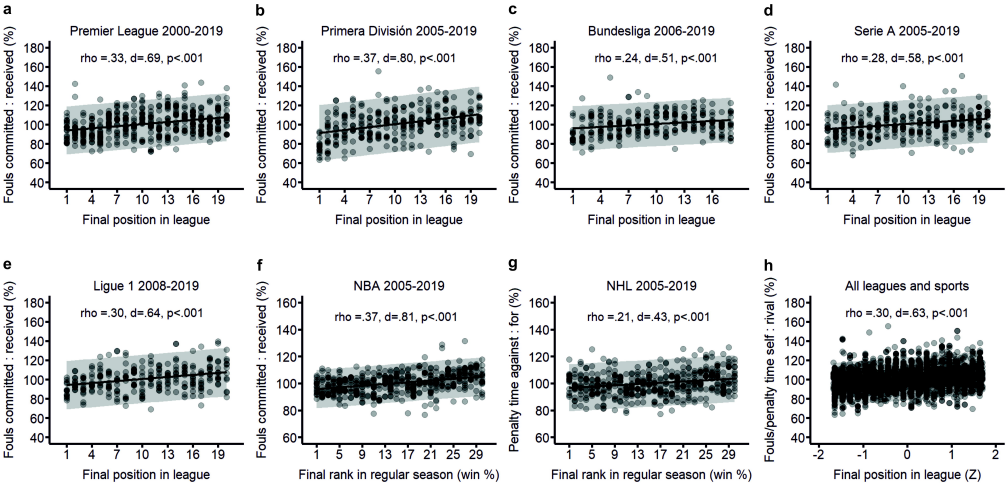
We subsequently sought for associations between neural reactivity to status signals and aggression on a within-subject basis. We did so by comparing the parametric modulator for aggression against the high- versus low-status rival using a paired t-test. This analysis tests for brain areas showing transient increases in activity as a function of participants’ trial-wise aggression levels against one opponent relative to the other. Five participants had to be excluded from this analysis due to a lack of variability in behavior (i.e. they chose the same punishment in all trials). We applied a threshold of $p<.001$ at the voxel level with a $p<.05$ family-wise error (FWE) correction at the cluster level.

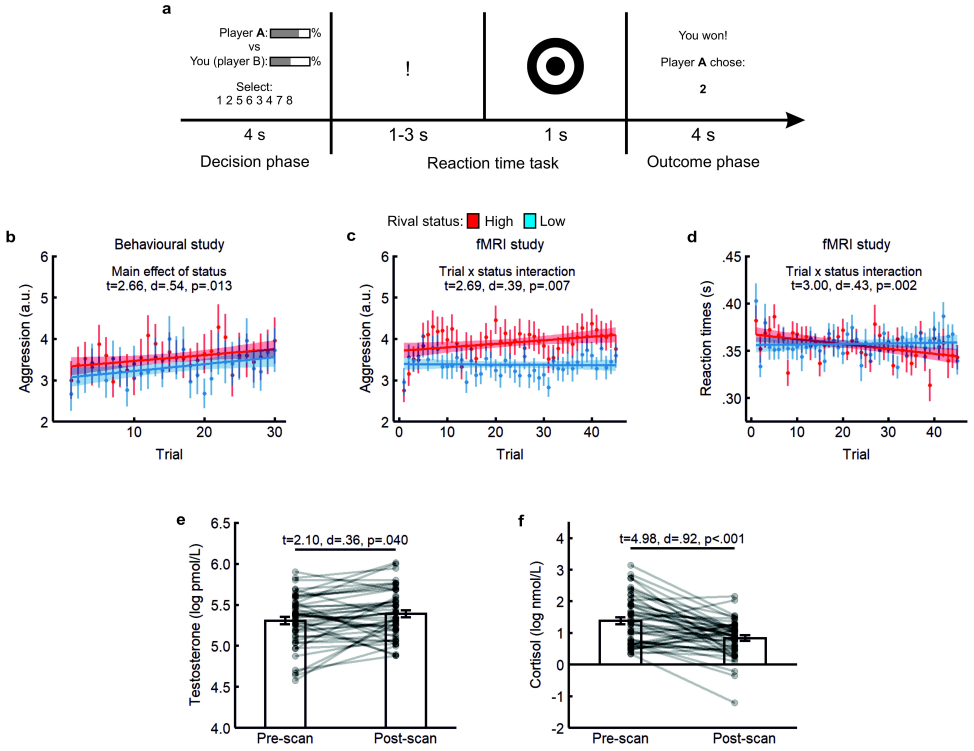
In the outcome phase, we tested whether wins and losses were differentially processed depending on the opponent’s status using a flexible factorial analysis. First, we tested the main effects of won vs lost in order to inspect whether we could reproduce the ventral striatum activity commonly observed during competitive victories (Table S1 in Supplementary Material). Because subjects lost more often against one of the two opponents, status and outcome were not orthogonal. We therefore did not inspect the main effect of status. Rather, we directly tested the interaction between the two factors, namely the contrast *[won high + lost low] > [lost high + won low]* and its opposite using paired t-tests (see

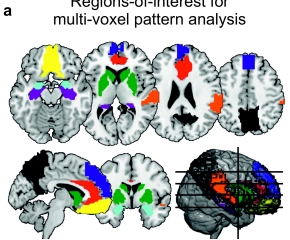
Table S1 in Supplementary Material for complete results). For these contrasts we also set a voxel-level $p<.001$ threshold with a cluster-wise $pFWE<.05$ correction.

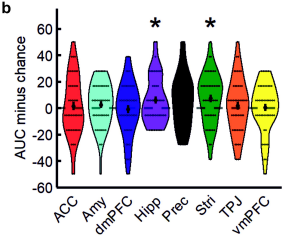
Finally, we probed whether interindividual differences in status-based aggression were associated with the neural processing of status signals. We used MarsBaR (http://marsbar.sourceforge.net/) to extract subject-wise parameter estimates from ROIs showing above-chance classification in the MVPA analyses (whole ROI) and from clusters showing significant activity in the outcome by status interaction (6 mm sphere around peak coordinate). We computed the correlation coefficients between activity in each ROI/cluster and mean aggression against the high- minus the low-status opponent. We also computed the 95% bootstrapped confidence intervals for the correlation coefficients of significant associations ($p<.05$) in order to assess their robustness and to protect against the influence of outlying observations. We performed all bootstrapping analyses using the *bootstrap* package (version 2019.5) in R (version 3.6.1) running on R Studio (version 1.1.423).

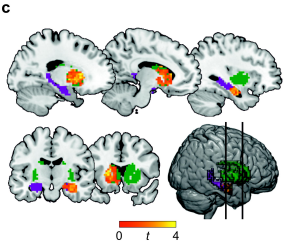
The neuroimaging study had also been approved by the Ethics Committee of the University of Lübeck. Subjects provided informed consent for participation and anonymized dissemination of the data. Again, we debriefed them regarding the goals and methods of the study. Behavioral and ROI data along with analysis scripts for this study are freely available via the Open Science Foundation (https://osf.io/2jvx4/). Raw and pre-processed neuroimaging data are available upon request. Our reporting of neuroimaging analysis and results adheres to best-practice recommendations [70].

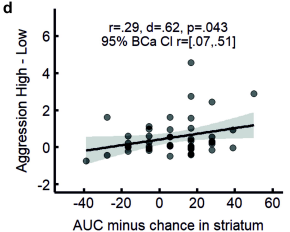

a
Premier League 2000-2019
rho =.33, d=.69, p<.001
Fouls committed : received (%)
Final position in league
b
Primera División 2005-2019
rho =.37, d=.80, p<.001
Fouls committed : received (%)
Final position in league
c
Bundesliga 2006-2019
rho =.24, d=.51, p<.001
Fouls committed : received (%)
Final position in league
d
Serie A 2005-2019
rho =.28, d=.58, p<.001
Fouls committed : received (%)
Final position in league
e
Ligue 1 2008-2019
rho =.30, d=.64, p<.001
Fouls committed : received (%)
Final position in league
f
NBA 2005-2019
rho =.37, d=.81, p<.001
Fouls committed : received (%)
Final rank in regular season (win %)
g
NHL 2005-2019
rho =.21, d=.43, p<.001
Penalty time against : for (%)
Final rank in regular season (win %)
h
All leagues and sports
rho =.30, d=.63, p<.001
Fouls/penalty time self : rival (%)
Final position in league (Z)

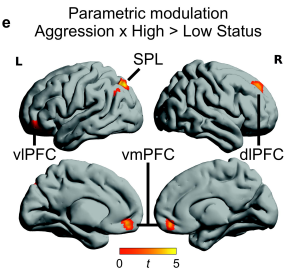

a
Player A: %
vs
You (player B): %
Select:
1 2 5 6 3 4 7 8
!
You won!
Player A chose:
2
4 s
1-3 s
1 s
4 s
Decision phase
Reaction time task
Outcome phase
Rival status: High Low
b
Behavioural study
Main effect of status
t=2.66, d=.54, p=.013
Aggression (a.u.)
Trial
c
fMRI study
Trial x status interaction
t=2.69, d=.39, p=.007
Aggression (a.u.)
Trial
d
fMRI study
Trial x status interaction
t=3.00, d=.43, p=.002
Reaction times (s)
Trial
e
t=2.10, d=.36, p=.040
Testosterone (log pmol/L)
Pre-scan
Post-scan
f
t=4.98, d=.92, p<.001
Cortisol (log nmol/L)
Pre-scan
Post-scan

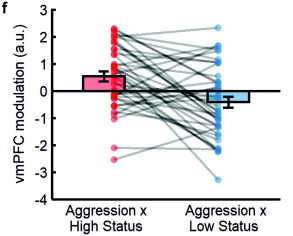

a
Regions-of-interest for
multi-voxel pattern analysis

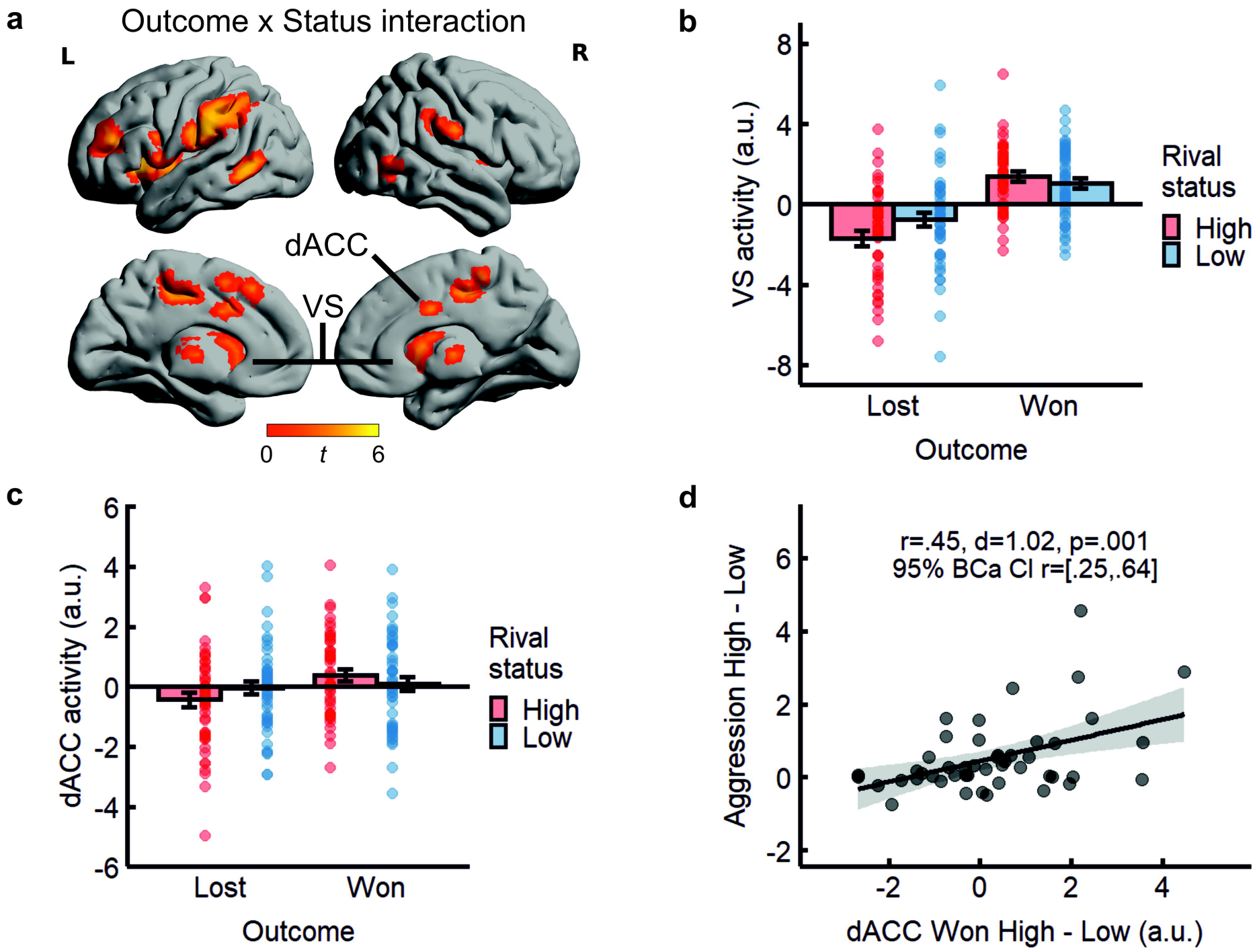

b
AUC minus chance
60
40
20
0
-20
-40
-60
*
*
ACC
Amy
dmPFC
Hipp
Prec
Stri
TPJ
vmPFC


c
0
t
4


d
r=.29, d=.62, p=.043
95% BCa CI r=[.07,.51]
Aggression High - Low
6
4
2
0
-2
-40
-20
0
20
40
60
AUC minus chance in striatum


e
Parametric modulation
Aggression x High > Low Status
L
R
SPL
vlPFC
vmPFC
dlPFC
0
t
5


f
vmPFC modulation (a.u.)
3
2
1
0
-1
-2
-3
-4
Aggression x
High Status
Aggression x
Low Status


a
Outcome x Status interaction
L
R
dACC
VS
0
t
6
b
VS activity (a.u.)
8
4
0
-4
-8
Lost
Won
Outcome
Rival status
High
Low
c
dACC activity (a.u.)
6
4
2
0
-2
-4
-6
Lost
Won
Outcome
Rival status
High
Low
d
r=.45, d=1.02, p=.001
95% BCa CI r=[.25,.64]
Aggression High - Low
6
4
2
0
-2
-2
0
2
4
dACC Won High - Low (a.u.)