Endogenous testosterone is associated with increased striatal response to audience effects during prosocial choices

Yansong Li, Elise Météreau, Ignacio Obeso, Luigi Butera, Marie Claire Villeval, Jean-Claude Dreher

Endogenous testosterone is associated with increased striatal response to audience effects during prosocial choices:

Yansong Li[a,b#], Elise Météreau[c,d], Ignacio Obeso[c,d,e], Luigi Butera[f], Marie Claire Villeval[g] and Jean-Claude Dreher[c,d#]

[a] Reward, Competition and Social Neuroscience Lab, Department of Psychology, School of Social and Behavioral Sciences, Nanjing University, Nanjing 210023, China
[b] Institute for Brain Sciences, Nanjing University, Nanjing 210023, China
[c] Neuroeconomics, Reward and decision making' group, Cognitive Neuroscience Center, CNRS UMR 5229, Bron, France
[d] Université Claude Bernard Lyon 1, Lyon, France
[e] HM Hospitales – Centro Integral en Neurociencias HM CINAC, Móstoles, Madrid, Spain
[f] Copenhagen Business School, Denmark
[g] Univ Lyon, CNRS, GATE (UMR 5824), Ecully 69130, France

Running Title: Testosterone and striatal response to audience effects

[#]Correspondence:

Yansong Li, Ph.D
Reward, Competition and Social Neuroscience Lab
Department of Psychology & Institute for Brain Sciences
Nanjing University
Nanjing, China
Telephone: 0086-025-89680958
Fax: 0086-025-89680950
Email: yansongli@nju.edu.cn

Jean-Claude Dreher, Ph.D
Neuroeconomics Lab,
Institut des Sciences Cognitives Marc Jeannerod, CNRS
67 Bd Pinel, 69675 Lyon
France
Tel: +33 (0)4 37 91 12 38
Fax:+33 (0)4 37 91 12 10
Email: dreher@isc.cnrs.fr

Abstract

The role of testosterone on cognitive functions in humans remains controversial. One recent hypothesis suggests that this steroid hormone advances social status. As being observed by others is known to modulate a range of behaviors because of image concerns, we hypothesized that such an audience effect might be an important component of status seeking that is under the control of testosterone. Thus, we investigated to which extent testosterone levels are associated with the effect of being observed during prosocial choices and the neural mechanisms underlying this effect. We enrolled twenty-four male participants, aged 22.47±2.62 years, in an fMRI experiment to examine the relationship between testosterone levels and brain activity engaged in deciding whether to accept or reject monetary transfers to two types of organizations (a positively evaluated organization and a negatively evaluated organization) in presence or absence of an audience. When comparing the public to the private condition, the rate of acceptance increased for the positively evaluated organization, while the rate of rejection increased for the negatively evaluated one. Higher testosterone levels were linked to greater activation in the striatum in the public compared to the private condition, regardless of the organization type. These results indicate a relationship between testosterone levels and striatal activity induced by the audience effect. These findings provide new insights on the role of testosterone in human social behavior.

Keywords: Testosterone; Audience effect; Striatum; Social image; Charitable giving

1. Introduction

The steroid hormone testosterone has long been known to regulate the development of physical masculinization (Renfree et al., 2002). Apart from its role in the body, there has been growing interest in understanding testosterone-behavior relationships over the past decades (Geniole and Carré, 2018; Hines, 2017). One traditional view on testosterone functions is that it drives certain forms of aggression in both humans (Coccaro et al., 2007; Dabbs and Hargrove, 1997; Räsänen et al., 1999) and non-human primates (Bouissou, 1983; Giammanco et al., 2005). However, this traditional view of the role of testosterone in driving aggression has been revisited in more recent theories and experiments (Archer, 2006; Nadler et al., 2019). Recent studies emphasized its relation to status-enhancing behavior in the form of prosocial or antisocial behavior, depending on the social contexts (Booth et al., 2006; Dreher et al., 2016; Eisenegger et al., 2011; Mazur and Booth, 1998). For example, higher levels of testosterone in both men and women have been associated with enhanced social status (Rowe et al., 2004; Sellers, 2006) or increased spatial cognitive skills when status is at play (Newman et al., 2005)**.** Other behavioral results in men and women have also emphasized the relationship between testosterone levels and social cooperation (Casto and Edwards, 2016; Sanchez-Pages and Turiegano, 2010) or the choice of an interaction strategy (domination vs. submission) in a social context (Inoue et al., 2017; van Honk et al., 2014). In addition to these correlational evidence, recent behavioral studies tested to what extent testosterone administration plays a causal role during social interactions. A single dose of testosterone in women decreased trust but increased generosity in non-competitive settings (Boksem et al., 2013)**,** led to fair bargaining behavior (Eisenegger et al., 2010) and motivated for reputable-status seeking, even when the resulting behaviors were economically disadvantageous (van Honk et al., 2016). Similarly, these findings have been extended to men. For example, exogenous testosterone administration in men has been

shown to increase not only the altruistic punishment of unfair offers, but also prosocial behavior (positive reciprocity) in response to generous offers in a modified ultimatum game (Dreher et al., 2016), social cooperation (van Honk et al., 2012), preferences for high-status goods (Nave et al., 2018) and status-seeking motivation with unstable low social status (Losecaat Vermeer et al., 2020).

However, a key element of social interactions in real-world settings is whether other individuals can observe both the decisions made by the decision maker and their consequences, which is in fact a neglected aspect of the aforementioned studies. Decisions under observability can indeed be influenced by individuals' image concerns. In these settings, individuals may focus on matching their in-group social values rather than raising social status (Everett et al., 2015). Previous studies have found that individuals' behavior can be influenced by the mere presence of others (Hamilton and Lind, 2016), suggesting that the presence of an audience may be one of the dominant factors driving several social enhancing behaviors (Bradley et al., 2018). Audience as a modulator of behavior has been found in a diversity of species, including humans and nonhuman primates (Chib et al., 2018; Sekiguchi and Hata, 2018). Given that the mere presence of an audience can promote status-seeking behavior in our social life and testosterone has been shown to play an important role in status-relevant behavior, understanding the extent to which testosterone levels can be related to audience during prosocial decisions would greatly advance our understanding of testosterone-behavior relationships. In particular, since testosterone is involved in status-relevant behavior, one may expect that an audience should enhance its relation with norm-compliant prosocial behavior. Moreover, identifying the underlying neural mechanisms of the association between testosterone levels and audience in prosocial behavior would provide important insights not only into the prosocial role of testosterone in the context of social interactions, but also into the mechanisms underlying the

testosterone-status relationship. This matters particularly since testosterone has been shown to be disrupted in psychiatric disorders (Li et al., 2020). In particular, children who have been exposed to high concentrations of testosterone as a fetus would be more likely to exhibit autistic traits (Mullard, 2009). Although previous research investigated the effect of audience on prosocial behavior in autism, the relationship with testosterone remains to be investigated (Izuma et al., 2011). Prior neuroimaging evidence pinpoints a brain network essential for conducting prosocial decisions. This includes the striatum, anterior cingulate cortex (ACC), ventromedial prefrontal cortex (vmPFC) and temporo-parietal junction (TPJ). This network is recruited when expecting social rewards as well as when weighing monetary costs against compliance with one's moral values, or when helping choices are made (Cutler and Campbell-Meiklejohn, 2019; Qu et al., 2019). Yet, how the aforementioned network is regulated by sex hormone is unknown, although an increasing effort has been devoted to exploring how other hormones such as estrogen and oxytocin modulate prosocial behavior (Kemp and Guastella, 2010; Zethraeus et al., 2009). Here, we explored the relationship between endogenous testosterone levels and the neural mechanisms underlying prosocial behavior in reaction to the presence or absence of an audience. To address this question, we used the behavioral data from a donation experiment published by Qu et al. (2019). In this experiment, participants had to decide whether to accept or reject monetary transfers to two organizations (one positively evaluated, and the other negatively evaluated). Prosocial behavior was characterized by two types of decisions: accepting a monetary transfer to a positively evaluated organization at a personal cost, or foregoing personal monetary gains to reject a transfer to an organization that they evaluated negatively. These decisions were made in private or in public, depending on the trials. Decisions while being observed required weighing the costs and benefits of accepting vs. rejecting the donation, plus the expected (positive or negative) image sent to the observer. Such reasoning requests the

conversion of social and monetary rewards into a common currency for comparisons to be made (Sescousse et al., 2015). In such settings, participants thus faced a moral dilemma: either serving a good cause but at a personal monetary cost, or making money but betraying ones' moral values. This design allows us to investigate whether testosterone is involved in guiding prosocial vs. selfish decisions induced by the presence of an audience when participants face a moral dilemma. Because weighing monetary costs against compliance with one's moral values (Qu et al., 2019) and perceiving one's good reputation (Izuma et al., 2008, 2010) have been reported to result in striatal activity, we hypothesize a positive correlation between testosterone levels and striatal activation while making prosocial decisions in reaction to the presence of an audience.

2. Material and methods

2.1 Participants

We summarize in this section the experimental design, all the details being developed in Qu *et al*. (2019). Twenty-four healthy male participants, aged 22.47±2.62 years, with no history of neurological or psychiatric illness participated in the fMRI experiment. Three participants were discarded from the analysis because of failure to collect testosterone data. All participants were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), and presented no symptoms of depression, as assessed by the 13-item version of the Beck Depression Inventory (Beck and Beck, 1972). Informed consent was obtained from every participant. The study was approved by the local ethics committee (CPP Centre Léon Bérard).

2.2 Pre-testing

As described in our previous study (Qu et al., 2019), a behavioral pilot study involving 48 healthy volunteers was performed at GATE-Lab, Lyon, to help with designing stimuli and task procedures. To guide the selection of the organizations, we asked them to complete a questionnaire after the presentation of brief descriptions and logo images of 14 organizations. Organizations with positive or negative valence were presented. For each one, participants had to rate their feelings towards them on a scale from −10 to 10. The organizations were presented in the questionnaire in a random order across participants. Based on this pilot study, we chose for the fMRI experiment the two organizations that received the worst (mean = −5.73, SD = 3.68) and the best (mean = 8.40, SD = 2.04) ratings. They were a negatively evaluated organization (NEG ORG) (‘Groupe d’Action Royaliste’, – an organization that aims at promoting the restoration of monarchy in France) and a positively evaluated charity (POS ORG) (‘Resto du coeur’, a charity providing food to poor people). Because the policy does not allow us to publish trademarked names, we have

changed the real names of these two organizations. GAR represents the NEG ORG and RES (a symbol of heart) represents the POS ORG (a charity providing food to poor people) (**Fig 1**).

2.3 Experimental Task

Our previous study (Qu et al., 2019) described that "we used a 2 × 2 within-participant design, in which participants decided whether to accept or reject monetary transfers to the two organizations. Depending on the blocks of decisions, the offers of transfer is concerned with either the POS ORG or the NEG ORG. Decisions were made either in presence or absence of observers ("public" vs. "private" conditions) (**Fig 1**). At the beginning of the experiment, participants received an initial endowment of 14 Euros. During the experiment, they were faced with successive offers involving a variable monetary payoff for themselves and a variable payoff for the organization. When making decisions regarding the POS ORG, participants had to decide whether to accept or reject monetary transfers to the organization at a variable monetary cost to themselves, deducted from their initial endowment. When making decisions regarding the NEG ORG, they had to decide whether to accept or reject monetary transfers to the organization in exchange for a personal monetary payoff added to their initial endowment. In the latter case, the only way for a participant to earn money was to accept a donation to the NEG ORG, whereas in the former treatment, any donation to the POS ORG involved a monetary loss for the participant. One crucial aspect is that in both treatments, each organization would receive a donation; however, in one case such a donation entails a moral cost for the individual (allowing the experimenter to send money to the NEG ORG in order to earn money for oneself may violate one's moral values), while in the other case, the donation to the organization generates a moral benefit for the individual (altruistically foregoing a personal gain to benefit the POS ORG may comply with one's moral values). Because we systematically varied the monetary cost of a moral decision, we

were able to identify the price elasticity of demand for moral actions. Intuitively, if participants did not perceive some actions as immoral, they would display no elasticity to the moral cost of choosing the self-serving action. The monetary stakes for the organizations and for the participants varied independently across trials. In each trial, the organization's potential gains ranged from 4 to 32 Euros, in increments of 4 Euros. Participants' potential payoffs (in the case of the NEG ORG) or costs (in the case of the POS ORG) varied from 1 to 8 Euros, in increments of 1 Euro. Each participant was therefore exposed to 64 different dilemmas.

Only one public decision and one private decision among all the trials were randomly selected for payment at the end of the experiment. If the participant accepted the offer in the randomly selected trial, the amount of the accepted transfer was sent to the organization (the mean of the two amounts was used if the two trials concerned the same organization), and the participant's endowment was increased or decreased based on his decision. If the same organization happened to be randomly selected twice, then the organization received the average transfer and the participant's endowment was adjusted based on the average of the two decisions. If the participant rejected the offer in the randomly selected trials, nothing was sent to the organization, and the participant's initial endowment was not modified.

The presence or absence of an observer (public versus private conditions) was displayed on the screen in the following way. In private trials, a yellow frame surrounded the offer, and a picture of a padlock was displayed at the top of the screen reminding participants about the privacy of their decisions. In the public condition, a cyan frame surrounded the offer, and a picture of the eyes of an observer was displayed above, reminding participants that an independent observer would see their decisions. Indeed, cues of being watched exert an influence on participants' behavior (Bateson et al., 2006). To further stress the visibility of their choices in the public trials, participants knew that an observer in the control room, to whom they were introduced prior to the experiment, would

see the participant's screen and therefore observe their public trials decisions; in the public trials, the chosen alternative was highlighted for 1.5 s on the screen by expanding the font, while the other option disappeared. In the private condition, no changes were made on the screen after the response, assuring participants that nobody would be able to see their choices from the scanner control room. Finally, at the end of the experiment, participants had to declare in front of a video camera which decision they made in the randomly selected trial for the public condition. Participants were told that decisions in the private condition were recorded anonymously, guaranteeing that none of the experimenters could link a participant's identity with his decisions. A person not affiliated with the experiment and unaware of its content paid all participants. All the participants reported believing in the manipulation.

For each possible combination of individual and organization payoffs, and for both organizations, participants made two decisions, one in private and one in public. Participants therefore made a total of 256 decisions, 128 related to the NEG ORG and 128 related to the POS ORG. Each trial began with the presentation of an offer, which could either be accepted or rejected by pressing the left or right button on a response pad. A fixation cross was displayed during a random time interval (jitters), drawn from a uniform distribution between 2.5 and 6.5s. Participants were encouraged to make their decision within 3 s. After this delay, a message was displayed on the screen to remind them to respond.

The scanning session was divided into 4 runs of 64 trials. The first two runs concerned one organization and the last two concerned the other organization. Within the first run of each organization, the first half of the trials was either public or private, with the opposite for the subsequent run. The order of the private/public conditions in the second run mirrored the order of these conditions in the first run. The order of presentation of the organizations and of public/ private conditions was balanced across participants. Thirty-two dilemmas from the

64 possible combinations were presented in each run and each private/public condition. To guarantee that the two pairs of runs of each organization were balanced with respect to the payoffs for the individual and the organization, we assigned to one run the set of dilemmas composed by the participant's odd potential payoffs and the 4, 12, 20, and 28 potential amounts for the organization, while the other run was assigned the 32 remaining dilemmas of the matrix. Within this criterion, the order of the 32 dilemmas was randomized.

Visual stimuli were back-projected on a screen located at the head of the scanner bed and presented to the participants through an adjustable mirror located above their head. The presentation of the stimuli was controlled by Presentation © software (Neurobehavioral Systems), which also recorded trigger pulses from the scanner signaling the beginning of each volume acquisition."

2.4 Procedures

During a first interview (the pilot pre-testing), participants were asked to rate their feelings toward each of 14 organizations on a scale ranging from −10 to 10. Based on this pilot study, we chose for the fMRI experiment the two organizations that received the worst and the best ratings. For the fMRI experiment, we selected only participants who rated the POS ORG with a score greater than 0 and the NEG ORG with a negative score. The day of the experiment, participants first received instructions about the experiment.

After receiving the instructions, participants did a few free practice trials of all conditions in the control room of the fMRI and were allowed to ask questions. After the practice session, participants were asked to read a description of the two organizations. Before entering the fMRI room, they met with the independent observer. After scanning, the participants were debriefed. Participants filled a post-experimental questionnaire asking whether they truly perceived the different trials as independent, whether they believed in the difference

between private and public conditions, and whether they thought that the presence of the observer had influenced their decisions.

2.5 Testosterone Measurements

In order to minimize the effect of circadian hormone rhythms, all sessions were conducted between 1:45 PM and 3:45 PM. Prior to and after the scanning session, blood samples were obtained to detect the levels of plasma testosterone for each participant. Plasma total testosterone was used for the assay and was measured by a solid-phase, competitive chemiluminescent enzyme immunoassay, IMMULITE 2000 (Diagnostic Products Corporation, Los Angeles, CA). Intra- and inter-assay coefficients of variation were 7.2% and 8.2%, respectively. Such an assay had an analytical sensitivity of 0.5 nmol/L. Corrections for incomplete recovery were made using 3H-labeled internal standards (Déchaud et al., 1981; Rinaldi et al., 2001; Sabot et al., 1985). Free testosterone would be more interesting to investigate, but we did not record sex-hormone binding globulin (SHBG) allowing to compute free testosterone values. In spite of this, the measurement of total testosterone has still been argued to be effective in exploring the potential link between testosterone levels and neuropsychological functions in humans (Hua et al., 2016). In order to control for other variables affecting testosterone levels, participants were asked to practice little physical exercise during the appointment day and to refrain from any caffeine-containing food or drinks and cigarettes from at least one hour before the experiment started.

2.6 Behavioral Analysis

We characterized accepted trials in the POS ORG and rejected trials in the NEG ORG as “prosocial selection”, as these two options permit to a positively evaluated charity to earn money or avoid that a negatively evaluated organization receives money at a personal direct

or indirect cost to the participants (either through a reduction of the initial endowment or through foregoing a potential gain). By contrast, the rejected trials in the POS ORG and the accepted trials in the NEG ORG were both characterized as “selfish selection” because these options increased or preserved the initial endowment. Our previous study (Qu et al., 2019) has reported in detail the relationships between the parameters of the tasks and participants’ decisions, identified by using random-effects logistic models for each organization. Therefore, here we only report a brief and updated analysis of the main findings after having excluded the three participants from our previous study for whom we failed collecting hormones. A repeated-measures ANOVA on prosocial choices was conducted, with audience condition (public vs. private) and organization type (POS vs. NEG ORG) as within-participants factors. This is followed by Wilcoxon signed-rank tests for post-hoc testing.

2.7 fMRI Data Acquisition

The details of the fMRI acquisition and analysis have been reported in Qu *et al*. (2019). fMRI data was acquired on a 1.5 Tesla Siemens MRI scanner. The scanning was divided into 4 sessions. Blood-oxygenation-level-dependent (BOLD) signal was measured with gradient echo T2* weighted echo-planar images (EPIs). Twenty-six interleaved slices parallel to the AC-PC line were acquired per volume (matrix 64*64, voxel size = 3.4*3.4*4 mm, TR=2500ms, TE=60ms). We used a manual shimming within a rectangular region including the orbitofrontal cortex and the basal ganglia to improve the local field homogeneity. A high-resolution T1-weighted structural scan was subsequently acquired for each participant (matrix 256 × 256 × 176; voxel size = 1 × 1 × 1 mm; TR = 1,970 ms; TE = 3.93 ms; flip angle = 15).

2.8 fMRI Pre-processing

Data were pre-processed and analyzed using the SPM8 software package (Wellcome Department of Imaging Neuroscience, London) implemented in Matlab 7.10 (Mathworks, Natick, MA). The first four functional volumes of each session were removed to allow the BOLD signal to reach a steady state. The remaining images were slice-timing corrected, spatially realigned and unwarped to correct for motion artifacts. Unwarping was performed based on phase maps calculated using the Fieldmap SPM toolbox. Then in order to suppress the residual fluctuations due to interpolation errors from large motions, we used the motion adjustment algorithm provided in the ArtRepair toolbox (Mazaika et al., 2009) after a smoothing with a 4 mm full width at half maximum (FWHM) Gaussian kernel. This method is an alternative to adding motion regressors to the design matrix. The scan artifacts were then detected and repaired using both global intensity and scan-to-scan movement with the Artifact Repair algorithm from the ArtRepair SPM toolbox.

For each participant, the structural image was co-registered to the mean functional image, segmented into white and gray matter, and the gray matter was normalized to a standard gray matter template distributed by SPM8. The transformation parameters estimated in this step were applied to all functional images. Functional images were then spatially smoothed with a 7 mm FWHM Gaussian kernel.

2.9 fMRI Data Analysis

As described in our previous study (Qu et al., 2019), at the single-participant level, statistical analyses were performed using a GLM in which all regressors were modeled as delta functions and convolved with a canonical hemodynamic response function (HRF). We applied a high-pass filter with a cut-off of 128 s to the time series to remove low-frequency noise and baseline drifts, and we used an AR(1) model plus white noise to correct for temporal autocorrelation. Estimations were done in an explicit grey matter mask based on the tissue probability map provided by SPM.

Since the current study aims at exploring the relationships between testosterone levels and brain activity involved in the audience effect, we need to describe the analysis of audience effects based on our previous study (Qu et al., 2019). Specifically, we focused on a number of brain regions, such as those associated with making prosocial choices in the charity condition and those engaged with an audience effect, regardless of organization types or choices. We attempted to build a model including 8 regressors of interest at the time of “offer onset” in separate conditions 2 (accepted trials vs. rejected trials) × 2 (private vs. public) × 2 (POS vs. NEG ORG). We included the size of the potential gain for the organization and the size of the potential gain or loss for the participant with two orthogonal parametric regressors. Because little is known about the brain networks engaged when being observed (i.e., in the public condition) compared to when making decisions in private, regardless of the choice made, we performed two contrasts to test for the main effects of audience and privacy: public > private, and private > public, regardless of the organization types and participants’ choices. Given our specific a priori region of interest, we used small volume correction (SVC) with a threshold of $P < 0.05$ (FWE corrected) based on our a priori region of interest. The SVC was performed using a sphere with 10mm radius centering around the coordinate of peak voxel in the left and right putamen (left: -16, 14, -10; right: 12, 10, -4) derived from a previous studies on audience effect (Izuma et al., 2010) and in the left and right caudate nucleus (x, y, z = -17, 6, 13 and x, y, z = 18, 6, 9) derived from a previous study where charitable donation was investigated (Moll et al., 2006). Please note that these original coordinates in the Talairach space were transformed into the corresponding coordinates in MNI space using GingerALE 2.3. Given that we ran four SVC tests restricted to a single region, we have used a Bonferroni-corrected threshold of 0.05/4 = 0.013, accounting for the number of SVC tests.

For the correlational analysis between testosterone levels and striatal activity induced by the public *vs*. private contrast for both organizations, we employed the averaged testosterone levels between those measured prior to and after the scanning session in a simple regression analysis. To illustrate the correlation between testosterone levels and the patterns of activation, percentage signal changes were extracted in the functional ROIs of interest (left caudate and left putamen) using the MarsBar toolbox (http://marsbar.sourceforge.net).

3. Results

3.1 Audience effects

Our ANOVA analysis showed that there was a significant main effect of organization type on prosocial choices ($F(1,20) = 7.50$, $p < 0.05$), whereas there was not a significant main effect of audience condition on prosocial choices ($F(1,20) = 0.01$, $p > 0.05$). Moreover, there was a significant interaction between them on prosocial choices ($F(1,20) = 8.79$, $p < 0.01$). A Wilcoxon signed-rank test showed in the POS ORG, participants accepted significantly more offers on average in the public (70%) as compared to the private condition (66%; Wilcoxon $|Z| = 2.81$, $p < 0.01$, $r = 0.43$). In contrast, in the NEG ORG, participants accepted significantly less offers on average in the public (43%) relative to the private condition (47%; Wilcoxon $|Z| = 2.30$, $p < 0.05$, $r = 0.35$) (**Fig 2A**). This was further confirmed by color-coded heatmaps of the probability of accepted donations for transfers to the POS ORG and the NEG ORG, respectively (**Fig 2B**). These color-coded heatmaps clearly demonstrated that participants were more willing to accept to donate to the POS ORG in public than in private condition and were less willing to accept to donate to the NEG ORG in public than in private condition.

3.2 The link between testosterone, behavior and striatum

We first analyzed the main effect of acceptance in the public compared to the private condition, independently of the organization type. Striatal activity significantly increased in public compared to private decisions (MNI [x y z] [-12 2 -2], T = 3.66, p(SVC) < 0.05, FWE) (**Table 1**). In addition, regions such as the anterior cingulate cortex (ACC) (MNI[x y z] [0 23 34], T = 5.65), temporal parietal junction (TPJ) (MNI[x y z] [48 -25 25], T = 4.34) were also active in public vs private decisions (**Table 1**). By contrast, in private vs public decisions, a different brain network was found with only the occipital gyrus (MNI[x y z] [30 -82 -20], T = 4.45) being significantly engaged (**Table 1**). Given that our a priori hypothesis predicts a

positive correlation between testosterone levels and striatal signal during prosocial decisions in presence of an audience, we performed a correlation analysis between testosterone levels and BOLD responses for prosocial decisions made in public vs in private for both types of organization. As predicted, our results revealed a positive relationship between testosterone levels and striatal activity induced by prosocial decisions for public > private condition (putamen: MNI[x y z] [-21 5 -11], T=6.77, *p*(SVC) < 0.05, FWE; caudate nucleus: MNI[x y z] [-15 2 13], T = 5.07, *p*(SVC) < 0.05, FWE) (**Fig 3; Table 2**). The striatum, involved in prosocial behavior in public, showed a correlation with endogenous testosterone levels. By contrast, when looking at prosocial decisions made in private > in public for both types of organization, we found no supra-threshold activations in the social image-related brain network correlating with testosterone levels (**Table 2**). Meanwhile, to exclude potential confounding effects caused by the salience of the public context per se, we have further performed correlational analyses between testosterone levels and striatal activities induced by prosocial decisions for public vs. implicit baseline and for private vs. implicit baseline. These analyses failed to reveal significant correlations between them (**supplementary Tables 1 and 2**). Moreover, considering that our previous study has revealed that the audience effect (public > private) for both prosocial choices and for selfish choices commonly engaged a common brain network including the striatum (Qu et al., 2019), our further correlation analysis revealed that similar results could also be observed for selfish decisions for public > private condition (**supplementary Figure 1**). This provides further evidence that such a relationship was not specific to prosocial decisions only but can also be observed for selfish choices. Taken together, these results somewhat indicate that our observation of a significant relationship between testosterone levels and audience-induced striatal activities was not driven by the salience of the public context per se. Finally, to further examine the potential relationship between testosterone levels and the difference in

prosocial decisions made in public vs. in private for each type of organization, we performed a number of correlational analyses. However, we did not observe any significant relationship between them (POS ORG: r = -0.07, p = 0.78; NEG ORG: r = 0.18, p = 0.44) (**supplementary Figure 2**). Similarly, when exploring the possible link between striatal activities in public vs. in private and the difference in prosocial decisions made in public vs. in private for each type of organization, we found a significant relationship between them neither for the POS ORG (putamen: r = 0.02, p = 0.94; caudate: r = -0.18, p = 0.43), nor for the NEG ORG (putamen: r = -0.16, p = 0.50; caudate: r = 0.03, p = 0.90).

4. Discussion

The goal of the present study was to investigate the relationship between endogenous testosterone levels and the neural correlates responsible for prosocial decisions in presence of an audience, *i.e.,* when social image and status concerns may be activated. Our results showed that striatal response correlated positively with endogenous testosterone levels in the public condition as compared to the private condition, regardless of organization types. That is, when being observed, a greater striatal activity correlated with testosterone levels. This effect highlights the fact that audience facilitates prosocial decisions for both types of organizations.

The striatum has been previously demonstrated to be one of the key brain areas involved in reputation-based behaviors, such as charitable giving and decision making in presence of a moral dilemma (Izuma, 2012; Izuma et al., 2010; Moll et al., 2006; Shenhav and Greene, 2010). This brain region is also strongly involved in reward processing (Haber and Knutson, 2009; Sescousse et al., 2013). To better understand the potential role of the striatum in public prosociality, two interdependent processes need to be considered during the process of reputation building. The first is the ability to create meta-representations of oneself so as to achieve the desirable image benefit from a given social behavior. A second process is the ability to overcome the conflict between the expected value of an option and the value of the other, less appealing, options (cost-benefit trade-off). While the right temporal parietal junction (TPJ) may contribute to each of these processes (Obeso et al., 2018), the striatum may preferentially be engaged in the cost-benefit analysis of the available options when image concerns are active (Izuma, 2012). This functional role of the striatum in reputation-based processes may be linked to the value attributed to rewards as a common denominator between prosocial behavior (monetary gains for the charity in the POS ORG) and moral behavior (moral benefit of rejecting offers in the NEG ORG). This was

probably the case for both organizations when choices were made in public rather than in private. In fact, Izuma *et al.* (2008, 2010) and Qu *et al*. (2019) have shown that making donations while being observed and receiving monetary rewards both elicit striatal regions activity. In addition, the striatum is known to be engaged upon recognition of acceptance from others, *i.e*., being liked by others (Davey et al., 2010). The results of the current study additionally reveal the neural mechanisms underlying the role of testosterone in public prosociality when facing different moral dilemmas.

One important question is to identify the exact processes underlying the relationship between testosterone levels and striatal activation. In our study, this process cannot be attributed to the standard role of testosterone in reactive aggression. Yet, testosterone levels have been shown to correspond with increased striatal activity related to monetary rewards (Op de Macks et al., 2011). Because striatal activity is engaged with different types of rewards (Li et al., 2015), including moral benefits, the observed correlation in the current study could be proposed to reflect that testosterone potentiates striatal circuits functionality to raise their reward functions, perhaps mediated through dopamine (Haber and Knutson, 2009). To sum up, our results could contribute to the understanding of the striatal functions in contexts where social image is at play, and reveal a further striatal role in social interactions, as testosterone levels might contribute to transform social image concerns into generous or prosocial acts even for individuals that are not intrinsically prosocially motivated.

Another possible contribution of the current findings to the literature relies on the translation from women's to men's prosocial behavior in public. Previous studies have shown the role of testosterone in status-enhancing behavior in women (Boksem et al., 2013; Eisenegger et al., 2010; Mehta et al., 2015; van Honk et al., 2012; Zilioli et al., 2014). However, it should be noted that these actual effects observed after testosterone treatment

were induced by factors other than testosterone, since in the female brain aromatization to estradiol could equally well mediate the behavioral effects. Also, testosterone administration induces supra-physiological levels that are not representative for the actual natural level of testosterone in the female brain. Our present study adds to the literature by showing how natural testosterone is related to striatal activity during prosocial behavior induced by the presence of an audience in men. This may suggest that the role of testosterone in social behavior could be observed across sexes. However, several cognitive functions have been proven different between men and women, as well as in temperament characteristics (Borkenau et al., 2012a; Eagly, 2013). Men seem to show more of a variable pattern of social characteristics than women such as in extraversion, or agreeableness levels, suggesting that women have a less variable personality across the general population (Borkenau et al., 2012b). These factors may induce sex differences in the interpretation of social contexts. For example, sex differences were reported with regard to cortisol levels disparity, which altered behavior differently in a competition context (Kivlighan et al., 2005). As such, this raises an interesting question of whether our current findings in men would extend to women.

Limitations

We acknowledge some limitations of our study. First, it included a relative small sample, possibly tempering the strength of our conclusions. Replications with larger samples would be welcome. Second, even though the present study had a strong prior hypothesis about the striatum involved in the audience effect (Izuma et al., 2010; Moll et al., 2006), it will still be useful to search for information about other brain regions since one region is unlikely to be working all by itself. Third, although the measurement of total testosterone has been argued to be effective in examining the relationship between testosterone and neuropsychological function (Hua et al., 2016), further correlation with free testosterone would be needed to

avoid limiting testing correlation to total testosterone levels, which may overlook the possibility of excessive bondage to either sex-hormone binding globulin (SHBG) or albumin in the blood. Fourth, we used blood samples to measure testosterone levels, which may have activated anticipatory stress leading to increased cortisol levels. Moreover, there is a great deal of interaction within the endocrine system, so our understanding of the relationship between testosterone levels and the audience effect on prosocial behavior would benefit from the inclusion of more hormones in the same study. In particular, the dual-hormone hypothesis posits that testosterone's role in status-motivated behavior is modulated by concentrations of cortisol (Dekkers et al., 2019; Mehta and Josephs, 2010; Mehta and Prasad, 2015). Due to the small sample size, we were not able to explore potential interacting effects of testosterone and cortisol on the audience effect on prosocial behavior. Fifth, the present study only concerns men. We chose to scan only men because gender has been shown to affect prosocial behavior (Buckholtz et al., 2015; Croson and Gneezy, 2009; FeldmanHall et al., 2015) and unethical behavior (Berns et al., 2012; Dreber and Johannesson, 2008; FeldmanHall et al., 2012). Moreover, young women experience hormonal modulations of the reward system (Andreoni and Vesterlund, 2001; Dreher et al., 2007), which may affect the testosterone levels. In addition, there are known interactions between the effects of audience and the observer’s gender (kept constant in the present experiment). For example, in women the mere presence of men can induce transient decrements in cognitive efficiency and academic performances when confronted to math tests despite similar performances when tested separately (Childs, 2012; Eckel and Grossman, 1998). There is no doubt that future studies should investigate whether the present findings extend to women. Sixth, although the present study provided novel insight on the relationship between testosterone levels and audience effect on prosocial behavior through striatal activity, it is only correlative evidence. Further investigations should explore

the causal role of testosterone on the audience effect, using exogenous testosterone administration.

5. Conclusion

The current study provides direct correlational neural evidence for prosocial image seeking in the striatum that is regulated by testosterone. These findings help with sheding light on prior findings showing that testosterone is involved in social status seeking in social endeavors. Our results constitute a good starting point for investigating the neural mechanisms underlying the causal role of testosterone in human social behaviors. Exploring the causal role of testosterone on the striatal activity induced by the audience effect by using exogenous testosterone administration would be a natural extension.

Acknowledgements

Y. L. was supported by the National Natural Science Foundation of China (Grant Number 31600929), the Fundamental Research Funds for the Central Universities (Grant Number 010914380002) and a grant from Society for Social Psychology of Jiangsu Province (Grant Number 20SSXGH006). This work has also benefited from the financial support of IDEXLYON from Université de Lyon (project INDEPTH) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005) and of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency. This work was further supported by grant from the Fondation pour la Recherche Médicale FRM DPA20140629796.

Author contributions

EM, MCV, LB and JCD contributed to the study concept and design. Testing and data collection were performed by EM. EM, IO and YL performed the data analysis. YL drafted the manuscript. IO, LB, MCV and JCD provided critical revisions of the manuscript for submission. All authors approved the final version of the manuscript for submission.

Declarations of interest: none.

References

Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. The Quarterly Journal of Economics 116, 293-312.

Archer, J., 2006. Testosterone and human aggression: an evaluation of the challenge hypothesis. Neuroscience and biobehavioral reviews 30, 319-345.

Bateson, M., Nettle, D., Roberts, G., 2006. Cues of being watched enhance cooperation in a real-world setting. Biol Lett 2, 412-414.

Beck, A.T., Beck, R.W., 1972. Screening depressed patients in family practice. A rapid technic. Postgraduate Medicine 52, 81-85.

Berns, G.S., Bell, E., Capra, C.M., Prietula, M.J., Moore, S., Anderson, B., Ginges, J., Atran, S., 2012. The price of your soul: neural evidence for the non-utilitarian representation of sacred values. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 754-762.

Boksem, M.A.S., Mehta, P.H., Van den Bergh, B., van Son, V., Trautmann, S.T., Roelofs, K., Smidts, A., Sanfey, A.G., 2013. Testosterone Inhibits Trust but Promotes Reciprocity. Psychological Science 24, 2306-2314.

Booth, A., Granger, D.A., Mazur, A., Kivlighan, K.T., 2006. Testosterone and social behavior. Social Forces 85, 167-191.

Borkenau, P., Hrebickova, M., Kuppens, P., Realo, A., Allik, J., 2012a. Sex Differences in Variability in Personality: A Study in Four Samples. Journal of personality.

Borkenau, P., McCrae, R.R., Terracciano, A., 2012b. Do men vary more than women in personality? A study in 51 cultures. Journal of research in personality.

Bouissou, M.F., 1983. Androgens, aggressive behaviour and social relationships in higher mammals. Hormone research 18, 43-61.

Bradley, A., Lawrence, C., Ferguson, E.J.P.o.t.R.S.B.B.S., 2018. Does observability affect prosociality? 285, 20180116.

Buckholtz, J.W., Martin, J.W., Treadway, M.T., Jan, K., Zald, D.H., Jones, O., Marois, R., 2015. From blame to punishment: disrupting prefrontal cortex activity reveals norm enforcement mechanisms. Neuron 87, 1369-1380.

Casto, K.V., Edwards, D.A., 2016. Testosterone and reconciliation among women: after-competition testosterone predicts prosocial attitudes towards opponents. Adaptive Human Behavior and Physiology 2, 220-233.

Chib, V.S., Adachi, R., O'doherty, J.P.J.S.c., neuroscience, a., 2018. Neural substrates of social facilitation effects on incentive-based performance. 13, 391-403.

Childs, J., 2012. Gender differences in lying. Economics Letters 114, 147-149.

Coccaro, E.F., Beresford, B., Minar, P., Kaskow, J., Geracioti, T., 2007. CSF testosterone: relationship to aggression, impulsivity, and venturesomeness in adult males with personality disorder. Journal of psychiatric research 41, 488-492.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic literature 47, 448-474.

Cutler, J., Campbell-Meiklejohn, D., 2019. A comparative fMRI meta-analysis of altruistic and strategic decisions to give. NeuroImage 184, 227-241.

Dabbs, J.M., Jr., Hargrove, M.F., 1997. Age, testosterone, and behavior among female prison inmates. Psychosom Med 59, 477-480.

Davey, C.G., Allen, N.B., Harrison, B.J., Dwyer, D.B., Yücel, M., 2010. Being liked activates primary reward and midline self‐related brain regions. Human brain mapping 31, 660-668.

Déchaud, H., Lejeune, H., Garoscio-Cholet, M., Mallein, R., Pugeat, M., 1981. Radioimmunoassay of testosterone not bound to sex-steroid-binding protein in plasma. Clinical chemistry 35, 1609-1614.

Dekkers, T.J., van Rentergem, J.A.A., Meijer, B., Popma, A., Wagemaker, E., 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? Neuroscience & Biobehavioral Reviews 96, 250-271.

Dreber, A., Johannesson, M., 2008. Gender differences in deception. Economics Letters 99, 197-199.

Dreher, J.-C., Dunne, S., Pazderska, A., Frodl, T., Nolan, J.J., O'Doherty, J.P., 2016. Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. Proceedings of the National Academy of Sciences 113, 11633.

Dreher, J.-C., Schmidt, P.J., Kohn, P., Furman, D., Rubinow, D., Berman, K.F., 2007. Menstrual cycle phase modulates reward-related neural function in women. Proc Natl Acad Sci U S A 104, 2465-2470.

Eagly, A.H., 2013. Sex differences in social behavior: A social-role interpretation. Psychology Press.

Eckel, C.C., Grossman, P.J., 1998. Are women less selfish than men?: Evidence from dictator experiments. The economic journal 108, 726-735.

Eisenegger, C., Haushofer, J., Fehr, E., 2011. The role of testosterone in social interaction. Trends Cogn Sci 15, 263-271.

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., Fehr, E., 2010. Prejudice and truth about the effect of testosterone on human bargaining behaviour. Nature 463, 356-359.

Everett, J.A.C., Faber, N.S., Crockett, M., 2015. Preferences and beliefs in ingroup favoritism. Front Behav Neurosci 9, 15-15.

FeldmanHall, O., Dalgleish, T., Evans, D., Mobbs, D., 2015. Empathic concern drives costly altruism. Neuroimage 105, 347-356.

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., Mobbs, D., 2012. Differential neural circuitry and self-interest in real vs hypothetical moral decisions. Soc Cogn Affect Neurosci 7, 743-751.

Geniole, S.N., Carré, J.M., 2018. Human social neuroendocrinology: Review of the rapid effects of testosterone. Hormones and Behavior 104, 192-205.

Giammanco, M., Tabacchi, G., Giammanco, S., Di Majo, D., La Guardia, M., 2005. Testosterone and aggressiveness. Medical science monitor : international medical journal of experimental and clinical research 11, RA136-145.

Haber, S.N., Knutson, B., 2009. The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology 35, 4-26.

Hamilton, A.F.d.C., Lind, F., 2016. Audience effects: what can they tell us about social neuroscience, theory of mind and autism? Culture and brain 4, 159-177.

Hines, M., 2017. Gonadal Hormones and Sexual Differentiation of Human Brain and Behavior.

Hua, J.T., Hildreth, K.L., Pelak, V.S., 2016. Effects of Testosterone Therapy on Cognitive Function in Aging: A Systematic Review. Cogn Behav Neurol 29, 122-138.

Inoue, Y., Takahashi, T., Burriss, R.P., Arai, S., Hasegawa, T., Yamagishi, T., Kiyonari, T., 2017. Testosterone promotes either dominance or submissiveness in the Ultimatum Game depending on players' social rank. Scientific Reports 7, 5335.

Izuma, K., 2012. The social neuroscience of reputation. Neuroscience research 72, 283-288.

Izuma, K., Matsumoto, K., Camerer, C.F., Adolphs, R., 2011. Insensitivity to social reputation in autism. Proc Natl Acad Sci U S A 108, 17302-17307.

Izuma, K., Saito, D.N., Sadato, N., 2008. Processing of social and monetary rewards in the human striatum. Neuron 58, 284-294.

Izuma, K., Saito, D.N., Sadato, N., 2010. Processing of the incentive for social approval in the ventral striatum during charitable donation. J Cogn Neurosci 22, 621-631.

Kemp, A.H., Guastella, A.J., 2010. Oxytocin: prosocial behavior, social salience, or approach-related behavior? Biological psychiatry 67, e33-e34.

Kivlighan, K.T., Granger, D.A., Booth, A., 2005. Gender differences in testosterone and cortisol response to competition. Psychoneuroendocrinology 30, 58-71.

Li, Y., Ramoz, N., Derrington, E., Dreher, J.-C., 2020. Hormonal responses in gambling versus alcohol abuse: A review of human studies. Progress in Neuro-Psychopharmacology and Biological Psychiatry 100, 109880.

Li, Y., Sescousse, G., Amiez, C., Dreher, J.C., 2015. Local morphology predicts functional organization of experienced value signals in the human orbitofrontal cortex. J Neurosci 35, 1648-1658.

Losecaat Vermeer, A.B., Krol, I., Gausterer, C., Wagner, B., Eisenegger, C., Lamm, C., 2020. Exogenous testosterone increases status-seeking motivation in men with unstable low social status. Psychoneuroendocrinology 113, 104552.

Mazaika, P., Hoeft, F., Glover, G., Reiss, A.L., 2009. Methods and Software for fMRI Analysis for Clinical Subjects. Human Brain Mapping.

Mazur, A., Booth, A., 1998. Testosterone and dominance in men. The Behavioral and brain sciences 21, 353-363; discussion 363-397.

Mehta, P.H., Josephs, R.A., 2010. Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. Hormones and behavior 58, 898-906.

Mehta, P.H., Prasad, S., 2015. The dual-hormone hypothesis: a brief review and future research agenda. Current Opinion in Behavioral Sciences 3, 163-168.

Mehta, P.H., van Son, V., Welker, K.M., Prasad, S., Sanfey, A.G., Smidts, A., Roelofs, K.J.P., 2015. Exogenous testosterone in women enhances and inhibits competitive decision-making depending on victory–defeat experience and trait dominance. 60, 224-236.

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., Grafman, J., 2006. Human fronto–mesolimbic networks guide decisions about charitable donation. Proceedings of the National Academy of Sciences 103, 15623-15628.

Mullard, A., 2009. What is the link between autism and testosterone? Nature.

Nadler, A., Camerer, C.F., Zava, D.T., Ortiz, T.L., Watson, N.V., Carré, J.M., Nave, G., 2019. Does testosterone impair men's cognitive empathy? Evidence from two large-scale randomized controlled trials. Proceedings of the Royal Society B: Biological Sciences 286, 20191062.

Nave, G., Nadler, A., Dubois, D., Zava, D., Camerer, C., Plassmann, H., 2018. Single-dose testosterone administration increases men's preference for status goods. Nat Commun 9, 2433.

Newman, M.L., Sellers, J.G., Josephs, R.A., 2005. Testosterone, cognition, and social status. Hormones and behavior 47, 205-211.

Obeso, I., Moisa, M., Ruff, C.C., Dreher, J.-C.J.e., 2018. A causal role for right temporo-parietal junction in signaling moral conflict. 7, e40671.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97-113.

Op de Macks, Z.A., Moor, B.G., Overgaauw, S., Güroğlu, B., Dahl, R.E., Crone, E.A., 2011. Testosterone levels correspond with increased ventral striatum activation in response to monetary rewards in adolescents. Developmental Cognitive Neuroscience 1, 506-516.

Qu, C., Météreau, E., Butera, L., Villeval, M.C., Dreher, J.-C., 2019. Neurocomputational mechanisms at play when weighing concerns for extrinsic rewards, moral values, and social image. PLOS Biology 17, e3000283.

Räsänen, P., Hakko, H., Visuri, S., Paanila, J., Kapanen, P., Suomela, T., Tiihonen, J., 1999. Serum testosterone levels, mental disorders and criminal behaviour. Acta psychiatrica scandinavica 99, 348-352.

Renfree, M.B., Wilson, J.D., Shaw, G., 2002. The hormonal control of sexual development, Novartis Foundation Symposium. Chichester; New York; John Wiley; 1999, pp. 136-156.

Rinaldi, S., Dechaud, H., Biessy, C., Morin-Raverot, V., Toniolo, P., Zeleniuch-Jacquotte, A., Akhmedkhanov, A., Shore, R.E., Secreto, G., Ciampi, A., Riboli, E., Kaaks, R., 2001. Reliability and validity of commercially available, direct radioimmunoassays for measurement of blood androgens and estrogens in postmenopausal women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 10, 757-765.

Rowe, R., Maughan, B., Worthman, C.M., Costello, E.J., Angold, A., 2004. Testosterone, antisocial behavior, and social dominance in boys: pubertal development and biosocial interaction. Biological psychiatry 55, 546-552.

Sabot, J.F., Deruaz, D., Dechaud, H., Bernard, P., Pinatel, H., 1985. Determination of plasma testosterone by mass fragmentography using [3,4–13C]testosterone as an internal standard. Journal of Chromatography 339, 233-242.

Sanchez-Pages, S., Turiegano, E., 2010. Testosterone, facial symmetry and cooperation in the prisoners' dilemma. Physiology & behavior 99, 355-361.

Sekiguchi, Y., Hata, T.J.B.P., 2018. Effects of the mere presence of conspecifics on the motor performance of rats: Higher speed and lower accuracy.

Sellers, J.G., 2006. Testosterone and status seeking.

748 Sescousse, G., Caldú, X., Segura, B., Dreher, J.-C., 2013. Processing of primary and
749 secondary rewards: A quantitative meta-analysis and review of human functional
750 neuroimaging studies. Neuroscience & Biobehavioral Reviews.

751 Sescousse, G., Li, Y., Dreher, J.-C., 2015. A common currency for the computation of
752 motivational values in the human striatum. Soc Cogn Affect Neurosci 10, 467-473.

753 Shenhav, A., Greene, J.D., 2010. Moral judgments recruit domain-general valuation
754 mechanisms to integrate representations of probability and magnitude. Neuron 67, 667-677.

755 van Honk, J., Bos, P.A., Terburg, D., 2014. Testosterone and dominance in humans:
756 behavioral and brain mechanisms, New frontiers in social neuroscience. Springer, pp.
757 201-214.

758 van Honk, J., Montoya, E.R., Bos, P.A., van Vugt, M., Terburg, D., 2012. New evidence on
759 testosterone and cooperation. Nature 485, E4.

760 van Honk, J., Will, G.-J., Terburg, D., Raub, W., Eisenegger, C., Buskens, V., 2016. Effects
761 of Testosterone Administration on Strategic Gambling in Poker Play. Scientific Reports 6,
762 18096.

763 Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B., Hirschberg, A.L.,
764 Johannesson, M., 2009. A randomized trial of the effect of estrogen and testosterone on
765 economic behavior. Proceedings of the National Academy of Sciences 106, 6535-6538.

766 Zilioli, S., Mehta, P.H., Watson, N.V., 2014. Losing the battle but winning the war: uncertain
767 outcomes reverse the usual effect of winning on testosterone. Biol Psychol 103, 54-62.
768

769

Figure Legends

Figure 1. Experimental design. We used a 2x2 within-participant design, in which participants decide to accept or reject the possibility of doing a costly good action for the benefit of a positively evaluated organization (POS ORG) or avoiding a bad one which would advantage both them and negatively evaluated organization (NEG ORG), either in presence or absence of an audience (PUBLIC vs. PRIVATE). The amounts of the potential transfers to the organizations and of the potential costs or payoffs to the participants were varied independently across trials. In each trial, the organization potential gains ranged from 4 to 32 Euros, by steps of 4 Euros. The participants' potential payoffs (in the case of the NEG ORG) or costs (in the case of the POS ORG) varied from 1 to 8 Euros, by steps of 1 Euro. This manipulation resulted in 64 different dilemmas. Each trial began with the presentation of an offer that the participant could either accept or reject by pressing the left button response or the right button response, respectively. To further stress the presence of observers during public trials, the chosen alternative was highlighted for 1.5s by expanding its characters, while the other was disappearing. On the opposite, in the private condition, no changes were shown after the response, ensuring the participant that nobody would be able to see their choice. A fixation cross was eventually displayed during a random time interval.

Figure 2. Behavioral results. (A) Decisions modulated by the presence of an audience. The participants' rate of acceptance was significantly increased when decisions were observed in public than in private for the POS ORG. Similarly, for the NEG ORG, participants were significantly more likely to reject the propositions in public than in private. The results indicated that participants made status seeking behavior due to the presence of an observer. POS ORG, positively evaluated organization; NEG ORG, negatively evaluated organization. ** $p < 0.01$, * $p < 0.05$. Error bars represent standard errors of the mean. **(B) Color-coded heatmaps of the probability of acceptance to donate for each dilemma of**

the 8x8 monetary/moral gain/loss matrix. Warmer colors indicate higher probability of acceptance, whereas colder colors indicate lower probability of acceptance. One heatmap is drawn for each type of organization and each audience condition.

Figure 3. The correlation between testosterone levels and striatal activity. Activation in the striatum (putamen: MNI[x y z] [-21 5 -11], T=6.77, *p*(SVC) < 0.05, FWE; caudate nucleus: MNI[x y z] [-15 2 13], T = 5.07, *p*(SVC) < 0.05, FWE) was positively correlated with testosterone levels, regardless of the types of organization. The scatter plots indicate that the striatum involved in decisions about transferring to the POS ORG and NEG ORG respectively in public is particularly prominent in high-testosterone men. POS ORG, positively evaluated organization; NEG ORG, negatively evaluated organization.

Table Legends

Table 1. Foci of activation relating to decisions made in public as compared to that made in private and vice versa. All reported foci are thresholded at $p < 0.001$ voxel-wise uncorrected with $p < 0.05$ FWE cluster-wise correction except for regions marked with the sign * which survived at a SVC corrected threshold of $p < 0.05$, FWE.

Table 2. Foci of activation relating to the correlation between brain activity induced by decisions made in public vs that made in private and the testosterone levels for both organizations. All reported foci are thresholded at $p < 0.001$ voxel-wise uncorrected with $p < 0.05$ FWE cluster-wise correction except for regions marked with the sign * which survived at a SVC corrected threshold of $p < 0.05$, FWE.

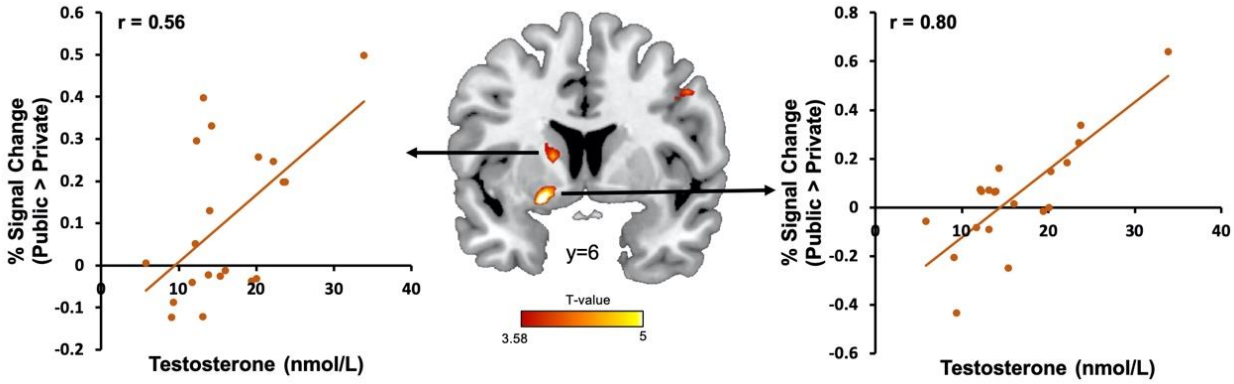
Supplemental materials:

Supplementary Table 1. Foci of activation relating to the correlation between brain activity induced by prosocial decisions for public *vs*. implicit baseline and the basal testosterone levels for both organizations. All reported foci are thresholded at p < 0.001 cluster-wise uncorrected (k > 10).

Brain regions	L/R	MNI coordinates			T value
		x	y	z	
Prosocial decisions: (public > implicit baseline) x testosterone levels					
Middle temporal gyrus	R	51	-1	-23	5.12
Superior parietal gyrus	L	-15	-34	37	4.63
Posterior cingulate gyrus	L	-34	43	26	4.31
Superior parietal gyrus	R	15	-40	58	4.43
Prosocial decisions: (implicit basline > public) x testosterone levels					
Lingual gyrus	L	-97	-5	64	4.81

Supplementary Table 2. Foci of activation relating to the correlation between brain activity induced by prosocial decisions for private *vs.* implicit baseline and the basal testosterone levels for both organizations. All reported foci are thresholded at $p < 0.001$ cluster-wise uncorrected ($k > 10$).

Brain regions	L/R	MNI coordinates			T value
		x	y	z	
Prosocial decisions: (private > implicit baseline) x testosterone levels					
Postcentral gyrus	R	51	-10	52	5.14
Anterior insula	R	33	-1	-8	5.03
Middle frontal gyrus	L	-42	14	52	4.98
Prosocial decisions: (implicit basline > private) x testosterone levels					
Lingual gyrus	L	-3	-82	-14	5.66
Cerebellum	L	-12	-37	-26	5.19
Midbrain	R	12	-19	-14	4.98



Supplementary Figure 1. The scatter plot showing the relationship between testosterone levels and striatal activity induced by selfish decisions made in public > private condition. Activation in the striatum (putamen: MNI[x y z] [-21 5 -11], T=6.77, *p*(SVC) < 0.05, FWE; caudate nucleus: MNI[x y z] [-15 2 13], T = 5.07, *p*(SVC) < 0.05, FWE) was positively correlated with testosterone levels, regardless of the types of organization. The scatter plots indicate that the striatum involved in selfish decisions about transferring to the POS ORG and NEG ORG respectively in public > in private is particularly prominent in high-testosterone men.