https://doi.org/10.1016/j.psyneuen.2023.106321

Running title: Testosterone and status-enhancing behavior

Testosterone administration modulates inequality aversion in healthy males: evidence from computational modeling

Jiajun Liao [1 #], Jianxin Ou [2 #], Yang Hu [3*], Philippe N. Tobler [4], Yin Wu [2, 5 *],

[1] School of Psychology, South China Normal University, Guangzhou, China

[2] Department of Applied Social Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong

[3] Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China

[4] Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Switzerland

[5] Research Institute for Sports Science and Technology, Hong Kong Polytechnic University, Hung Hom, Hong Kong

[#] These authors have contributed equally to this work.

[*] Author for correspondence: Dr. Yin Wu, Department of Applied Social Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong. Email: y.wu@polyu.edu.hk or Dr. Yang Hu, Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China. Email: yanghu@psy.ecnu.edu.cn

Abstract

Fairness concerns play a prominent role in promoting cooperation in human societies. Social preferences involving fairness concern have been associated with individual testosterone levels. However, the causal effects of testosterone administration on fairness-related decision making remain to be elucidated. Here, we used a randomized, double-blind, between-participant design and administered testosterone or placebo gel to 120 healthy young men. Three hours after administration, participants performed a modified Dictator Game from behavioral economics, in which they were asked to choose one of two monetary allocations between themselves and anonymous partners. Participants were either in a position of advantageous inequality (i.e., endowed with more than others) or disadvantageous inequality (i.e., endowed with less than others). Computational modeling showed that inequality-related preferences explained behavior better than competing models. Importantly, compared with the placebo group, the testosterone group showed significantly reduced aversion to advantageous inequality but enhanced aversion to disadvantageous inequality. These findings suggest that testosterone facilitates decisions that prioritize selfish economic motives over fairness concerns, which in turn may boost status-enhancing behaviors.

Keywords: advantageous inequality; disadvantageous inequality; inequality aversion; fairness-related decision-making; testosterone; computational model

1. Introduction

As one of the most other-regarding species, humans are interested in the relationship between their own payoffs and those of others and often show aversion to payoff inequality (Charness and Rabin, 2002; Fehr and Camerer, 2007; Fehr and Krajbich, 2014). Inequality can take two forms, depending on how goods are distributed. In advantageous inequality (AI), one receives more than the other person, whereas in disadvantageous inequality (DI), one receives less. Aversion to inequality implements a basic fairness norm for social preferences, which plays a key role in building and maintaining large-scale cooperative relationships, and thus serves as a cornerstone of human society (Fehr and Fischbacher, 2003; McAuliffe et al., 2017).

Behavioral neuroendocrinology has identified multiple associations between social preferences and hormones in human social decision-making. As one of the major sex steroids, testosterone plays a critical role not only in regulating physical development and reproductive behavior in both animals and humans (Gray et al., 2020; Muller, 2017), but also in modulating social cognition and behaviors (Eisenegger et al., 2011). One prominent line of research (Eisenegger et al., 2011) proposes that testosterone advances social status by promoting self-interested behaviors and keeping more resources for oneself. Studies have found that testosterone could facilitate status-enhancing behavior driven by selfish motives, such as competition for reward and resources (Geniole et al., 2019; Losecaat Vermeer et al., 2020), and reactive aggression (Geniole et al., 2019; Wagels et al., 2018, 2017). A separate body of work has shown that testosterone can increase prosocial behavior, which too may help individuals gain higher social status. For example, testosterone administration to men can lead them to reward

those obeying fairness norms (Dreher et al., 2016), and make them more willing to donate to charity when being watched by strangers (Wu et al., 2020). Thus, within the status-enhancing hypothesis, it remains unclear whether testosterone promotes status primarily through reduced or increased fairness concerns.

Previous research supports the notion that testosterone may regulate fairness-related decisions, at least in the Ultimatum Game (UG), a paradigm that measures the importance of fairness concerns during economic interactions between two individuals (Güth et al., 1982). For instance, Burnham (2007) and colleagues found that when acting as responders in the UG, participants with higher endogenous testosterone level were more likely to reject unequal offers. When acting as proposers, men who received exogenous testosterone were also more likely to make advantageously unequal offers than those receiving placebo (Zak et al., 2009). However, there exists conflicting evidence. For example, exogenous testosterone administration was found to increase fairness in women acting as proposers (Eisenegger et al., 2010). Another study found no reliable effect of exogenous testosterone on the rejection rate of unequal offers in men acting as responders (Dreher et al., 2016).

Although these studies have offered insights into the relationship between testosterone and fairness-related behaviors, two key issues remain poorly understood. First, few studies simultaneously investigated how testosterone altered the processing of advantageous and disadvantageous inequality. To our knowledge, the only study that has considered both contexts used the UG and did not find a reliable effect of testosterone on the probability of responders accepting offers that would create AI or DI (Dreher et al., 2016). Note though that responder decisions in the UG are influenced not only by

fairness considerations but also by other factors such as retaliation. Hence, it is not clear whether and how the effects of testosterone on fairness considerations vary across contexts.

Second, previous studies rarely examined the psychological mechanisms underlying the effect of testosterone on fairness-related decisions. While inequality aversion is an obvious underlying mechanism, it is possible that alternative factors contribute to fairness-related behaviors. For instance, people may simply consider the other's payoff (rather than inequality) or the total payoff of all parties (Charness and Rabin, 2002) during decision-making. Thus, formal model-based analyses are needed to rule out alternative interpretations and to reveal precisely how testosterone affects computation of value during fairness-related decision-making.

To address these issues, we conducted a study using a modified Dictator Game (DG) (Morishima et al., 2012) in combination with computational modeling in a double-blind, placebo-controlled, between-subject design. In this task, participants were asked to allocate money between themselves and anonymous partners. In each trial, they made binary choices between two unequal options. Critically, sometimes both options consistently earned participants more (AI context) or less (DI context) than the partner, whereas in other cases participants had to choose between advantageous and disadvantageous unequal options (mixed context). Note that in the DG, the monetary allocations are solely determined by the participants, which avoids potential confounds related to retaliation by the responder or strategic considerations by the proposer in the UG (counteracting possible retaliation by the responder) (Ruff et al., 2013). To further characterize the mechanisms by which testosterone modulated fairness-related decisions, we leveraged computational modelling approaches that allow

decomposing the psychological processes that underpin complex social behavior (Crockett, 2016; Konovalov et al., 2018). We specifically focused on the Fehr-Schmidt model (Fehr and Schmidt, 1999), a classic economic theory that captures aversion to advantageous and disadvantageous inequality. To rule out alternative mechanisms, we also compared the Fehr-Schmidt model with alternative models of social preference.

According to the status-enhancing account and previous findings mentioned above, we proposed two competing hypotheses regarding the effect of testosterone on inequality aversion. Hypothesis 1 expects testosterone to boost status-enhancing behavior by claiming more resources for oneself. In this view, participants with testosterone administration will exhibit decreased aversion to advantageous inequality and increased aversion to disadvantageous inequality. In other words, testosterone will reduce the disdain for earning more than others as well as increase the disdain for earning less than others. Hypothesis 2 proposes that testosterone promotes status-enhancing behavior as signaled by increased generosity, making exactly opposite predictions to Hypothesis 1. In this view, participants with testosterone administration will exhibit a stronger dislike of earning more than others and a stronger enjoyment of earning less than others.

2. Materials and methods

2.1 Participants

Using G*Power 3.1(Faul et al., 2007), we set effect size (Cohen's *d*) at 0.7, α at 0.05 and 1 – β at .95, resulting in a sample size of 110 participants (55 per group). One hundred and twenty healthy males were recruited for this study through university advertisement (Placebo group: $N = 60$, mean age =

20.84 years, *SD* = 2.14, age range = 18-26; Testosterone group: *N* = 60, mean age = 20.94 years, *SD* = 2.75, age range = 18-26). Only males were recruited because the dosing and pharmacokinetics of a single dose of Androgel® are established only for men (Eisenegger et al., 2013; Wu et al., 2018, 2019). We screened participants through a telephone interview and excluded individuals who were taking psychotropic medications or having any psychiatric/neurological disorders. Participants were instructed to abstain from alcohol, caffeine, and smoking for 24 h before the testing session. The study was conducted in accordance with the Declaration of Helsinki and was approved by the local research ethics committee. Written informed consent was obtained from all participants. Participants were compensated with 170 CNY (~ $25.5) as a participation fee and an additional payoff (average: 14.1 CNY, ~ $2.1) depending on their choices in the decision-making task (see below).

2.2 Testosterone administration

The study used a randomized, double-blind, placebo-controlled, between-participant design. All sessions started at 13:00 and lasted approximately 4 hours. Four participants unacquainted with each other came to the testing room individually around the same time and were assigned to four different cubicles for the decision-making task (see below). Participants in the testosterone group received a single dose of testosterone gel, containing 150 mg testosterone [Androgel®], while participants in the placebo group received a placebo gel. The testosterone and placebo gel as well as their packing were indistinguishable. In both groups, the gel was applied to the participant's shoulders and upper arms by a male research assistant, who was blind to both the experimental condition and the purpose of the study. The experimental task commenced three hours post-dosing in accordance with previous pharmacokinetic data (Carré et al., 2015; Eisenegger et al., 2013; Wu et al., 2018, 2019). During the

waiting period, participants were asked to stay in the testing room and were provided with newspapers and magazines that were unrelated to the present study.

2.3 Modified Dictator Game

The Dictator Game was modified from Morishima and colleagues (Morishima et al., 2012). Participants were informed that they would make choices regarding monetary rewards for themselves and another participant. On each trial, participants were first paired with a partner who was randomly selected from a previously collected data pool (1 ~ 2s). Next, the participant had 24 s to select one of two payoff allocations (i.e., option X and Y), which provided different amounts of money for the participant and the anonymous partner (see **Figure 1**). Once the participant made a choice, the selected option was highlighted by a yellow rectangle for 0.8 s.

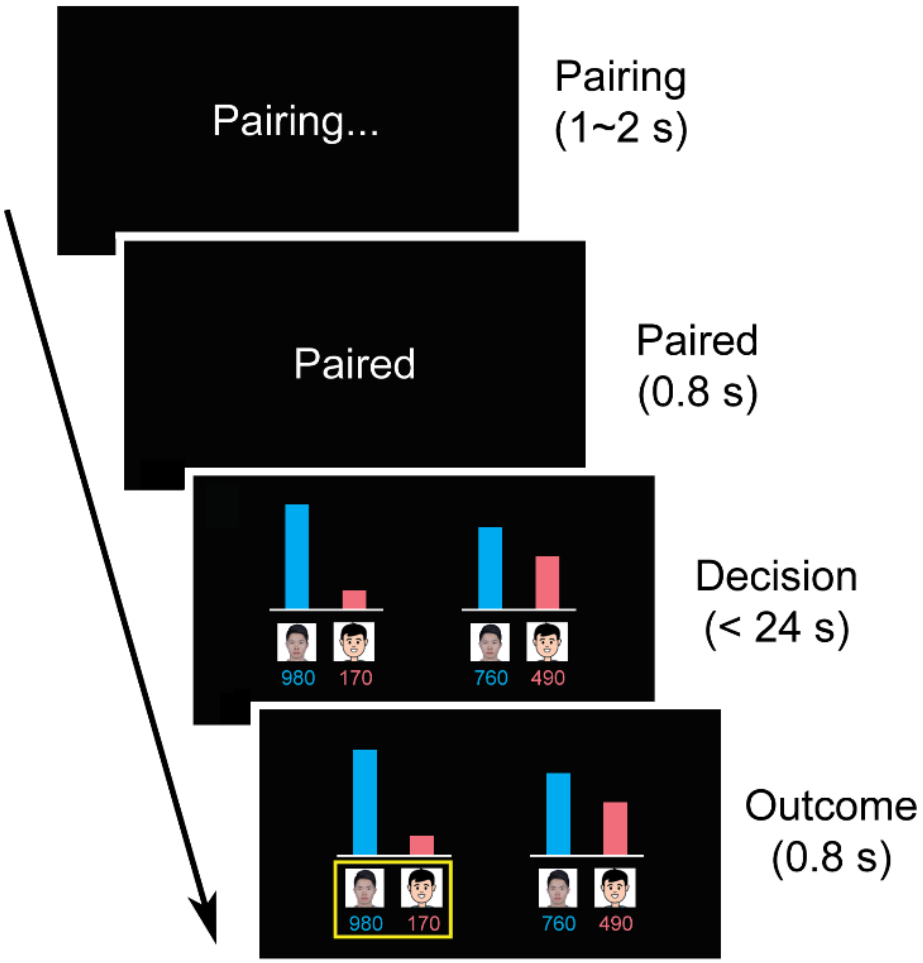

Figure 1. The Modified Dictator Game. Each trial began with a pairing period during which one interacting partner was randomly chosen. In the decision phase, participants were asked to choose between two options with different amounts of money for themselves and the anonymous partner. Once the decision was made, the chosen option was highlighted with a yellow rectangular. If no response was made within 24 s, this trial was aborted and the text "Pay attention" was presented on the screen for 2 s.

The task consisted of two identical blocks of 45 trials. Each block included 15 decisions in which the gains of the participant were larger than the gains of the partner for both option X and Y (AI context), and 15 trials in which the gains of the participant were smaller (DI context). The remaining 15 trials consisted of one option from the advantageous condition and one option from the disadvantageous condition (mixed context; see **Table S1** for details of options), which follows the study by Morishima *et al.* (2012). The positions of participant and partner benefits and the corresponding colors of frames

were counterbalanced across participants. Participants were informed that four trials would be randomly selected and implemented at the end of the experiment.

2.4 Data analyses

As the primary goal of this study is to investigate the computational mechanisms underlying the effects of testosterone during fairness-related decision-making, our analysis would mainly focus on the computational modeling of choice. We also performed additional model-free regression analyses to corroborate these model-based findings (see **Table S2** for summary of descriptive statistics). Note that a total of two trials in which participants did not respond within 24s were excluded from these analyses.

2.4.1 Computational Modeling of Choice.

We first assessed the Fehr-Schmidt model of inequality aversion. This model captures how strongly individuals are averse to advantageous inequality and disadvantage inequality. The utility function (equation 1) of each option is defined as follows:

$$U = M_{self} - \alpha \cdot \max(M_{other} - M_{self}, 0) - \beta \cdot \max(M_{self} - M_{other}, 0) \quad (1)$$

U denotes the utility (subjective value) of each option; M_{self} and M_{other} represent the amount of money for the participant and the anonymous partner, respectively. The parameters α and β measure preference towards disadvantageous or advantageous inequality respectively. Following a previous approach (Holper et al., 2018) we did not impose the constraints of the original model that α should be larger than, or equal to, β and that β should be positive. We additionally assessed a series of social preference models to rule out other mechanisms that could potentially explain choices (see **Table 1**).

These models were fitted to choices of all participants in a Hierarchical Bayesian approach (HBA; Gelman *et al.*, 2013) with Markov chain Monte Carlo (MCMC) sampling methods using the rstan packages in R (Carpenter et al., 2017). This model-fitting approach allowed us to estimate the individual- and group-level parameters simultaneously in a mutually constrained manner (Ahn et al., 2013). Model comparison was implemented with the Leave-One-Out Information Criterion (LOOIC; Vehtari *et al.*, 2017). Importantly, our hierarchical model consisted of two set of parameters, with one for the placebo group (group-level parameters: μ_i; individual-level parameters: i, where i refers to specific parameters in the model, same below) and the other for the between-group difference (i.e., a dummy variable characterizing the difference between testosterone and placebo; group-level parameters: $\Delta\mu_i$; individual-level parameters Δi). All of these parameters were estimated simultaneously using a hierarchical Bayesian approach, which provides full posterior distributions instead of point estimates.

For each individual-level parameter, we calculated the median of the posterior distribution for both groups as the best representative of point estimate. In particular, the individual-level parameter for the testosterone group was calculated by summing the median of the group-level parameter for the placebo group and that of the individual-level dummy parameter (i.e., $\mu_i + \Delta$i). Using these medians we performed the asymptotic two-sample Fisher-Pitman Permutation Test in R (Zeileis et al., 2008). We also report 95% highest density intervals (HDI) of all hyperparameters, i.e. the intervals which credibly covered most of the distribution for these parameters (Kruschke, 2014). A difference between the placebo and the testosterone groups was recognized as "significant" if the 95% HDI interval deviated from zero in dummy group variables (see **SOM** for details).

2.4.2 Model-Free Analyses.

We performed logistic linear mixed regression on choices using the *glmer* function in the "lme4" package (Bates et al., 2014). In particular, we examined how testosterone and inequality influences participants' choices by including *Group* (i.e., a dummy variable with reference level placebo) and *Difference of Inequality* (i.e., absolute difference in inequality between two options) as fixed-effect predictors. To account for the possibility that participants might base their decisions on their own payoffs, we also included *Difference of Self-payoff* (i.e., the absolute difference of payoff to oneself between two options) as a fixed-effect covariate. For random effects, we started with the simplest version, including only intercepts that varied across participants. Moreover, we tested other models which additionally included *Difference of Self-payoff*, *Difference of Inequality*, or both variables as random slopes. The best-fitting model was determined based on their Akaike Information Criterion (AIC) scores. In both AI and DI contexts, the choice was coded as 1 if the less unequal option was chosen and 0 otherwise. Note that trials in the mixed context were not used in the model-free analyses because any given choice in this context can reflect avoidance of one or appeal of the other type of inequality, making it challenging to interpret testosterone effects on behavior unequivocally. For completeness, we performed an exploratory analysis of choices in the mixed context and reported the results in the SOM (see Supplementary Table S5 for details).

3. Results

3.1 Computational Modelling of Choice

We first assessed which model explained decisions best. Bayesian model comparison indicated that the

Fehr-Schmidt (Fehr and Schmidt, 1999) model without constraints on α and β (Holper et al., 2018) outperformed competing models (see **Table 1**). Posterior predictive check showed that this model indeed captured the actual choices of participants in both groups (Pearson correlation: rs > 0.9, *p*s < .001; see **Figure S1**; also see **SOM** for details).

Next, we examined the effects of testosterone on the posterior means of the individual-level parameters in the best-fitting model. We found that aversion to advantageous inequality (i.e., β) was significantly smaller in the testosterone group than in the placebo group (mean ± SD: testosterone vs. placebo: 0.05 ± 0.48 vs. 0.25 ± 0.30; permutation-test, $Z = -2.52$, $p = .012$, Cohen's $d = 0.50$). Conversely, aversion to disadvantageous inequality (i.e., α) was significantly larger in the testosterone group than in the placebo group (mean ± SD: testosterone vs, placebo: 0.17 ± 0.70 vs. -0.03 ± 0.31; $Z = 1.97$, $p = .049$, Cohen's $d = 0.37$; see Figure 2). We did not observe a statistically significant difference in the inverse temperature parameter between the two groups (mean ± SD: testosterone vs, placebo: 0.04 ± 0.03 vs. 0.03 ± 0.03; $Z = 0.60$, $p = .551$, Cohen's $d = 0.33$), suggesting that testosterone had little effect on choice consistency. These results were also substantiated by analyses on the group-level parameters (see **Table S3** and **Figure S2**). Together, our data support Hypothesis 1 rather than Hypothesis 2 and suggest that testosterone boosts status-enhancing behavior by increased focus on selfish relative gains.

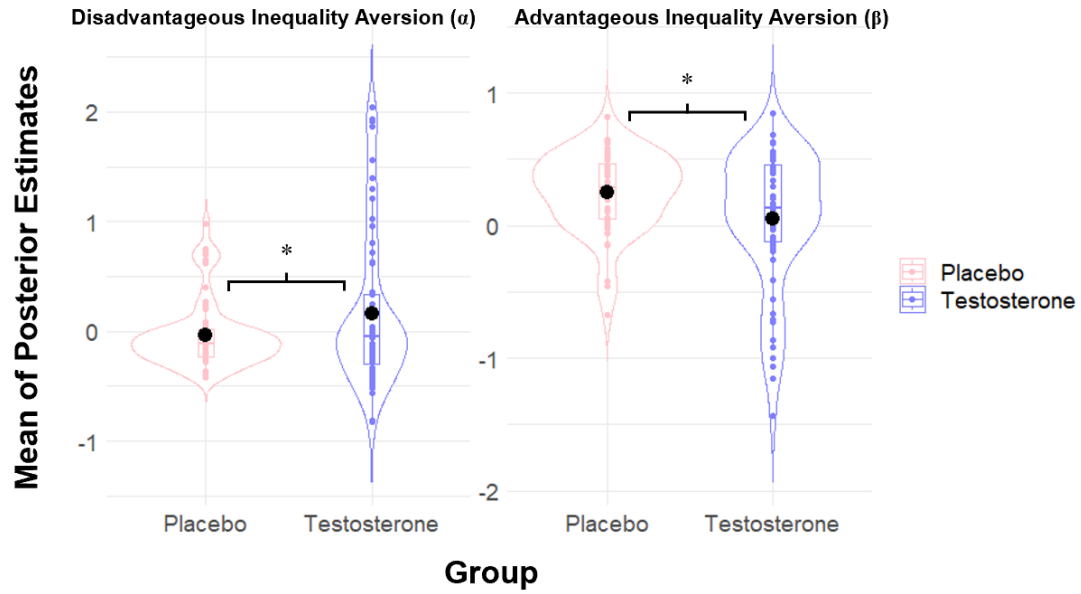

Figure 2. The individual-level parameter estimated from the best-fitting model (i.e., the unconstrained Fehr-Schmidt inequality aversion model). The curve in the violin plot stands for the distribution of the posterior median of individual-level parameters regarding inequality aversion (α and β) across all participants between two groups. The black dot indicates the mean of the posterior median of individual-level parameters for each group. The scattered dot indicates the data of each single participant. In the box plot, the line represents the median and the outer edges indicate the upper and lower quartiles. Significance: $^{*}p < 0.05$.

3.2 Model-Free Analyses

Note that the models with both variables (*Difference of Inequality* and *Difference of Self-Payoff*) as random slopes failed to converge. Among the remaining models, the one that included intercepts and Difference of Inequality as random slopes exhibited the lowest AIC scores in both AI and DI contexts, indicating the best-fitting model (see Supplementary **Table S4** for details). In the AI context, we found that choices were significantly affected by testosterone (vs. placebo; b = -1.05, SE = 0.38, Z = -2.74, p = .006). Specifically, individuals in the testosterone group were less likely to choose the less advantageous option, which is consistent with reduced aversion to advantageous inequality and our model-based results on β. In the DI context, individuals in the testosterone group tended to choose the

less unequal option more often compared to the placebo group (b = 1.01, SE = 0.56, Z = 1.80, p = .071). Although this difference was marginally significant, it aligns with increased aversion to disadvantageous inequality and our model-based results on α.

4. Discussion

The present study examined the effects of testosterone administration on fairness-related decision making and the underlying mechanisms combining an incentive economic task with computational modeling. Our computational model decomposed and elucidated the latent psychological processes underlying fairness-related decision-making and showed that testosterone affected aversion to both forms of inequality, indicating that testosterone increases relative selfishness. Together, we advance our understanding of the causal relationship between testosterone and fairness-related decisions and provide a novel perspective on the hormonal basis of social preferences.

A variant of the Fehr-Schmidt (1999) model (Holper et al., 2018) of inequality aversion better accounted for participants' choices than competing models with different underlying mechanisms. For example, overall (absolute) inequality explained decisions less well than the unconstrained Fehr-Schmidt model, indicating that the sign of the inequality matters. In addition, our winning model fitted the choice data better than the model hypothesizing the tradeoff between self-reward and other-reward, suggesting that payoff inequality, rather than only self-reward, is affected by testosterone. The Fehr-Schmidt model also outperforms the alternative model assuming the efficiency concern, which suggest that individuals do care the payoff inequality more than the total payoff during fairness-related decision making.

Converging with the model-free findings, participants in the testosterone group were more AI-seeking and DI-avoiding than participants in the placebo group. These findings support Hypothesis 1 and suggest that testosterone enhances status by boosting selfish rather than generous motives and behavior. According to the status-enhancing hypothesis, the key function of testosterone is to help individuals (esp., males) maintain, and improve upon, their current status (Carré et al., 2009; Carré and Olmstead, 2015). One method of achieving material status is to take more resources for oneself and thereby violating basic fairness norms. Indeed, previous evidence robustly associated testosterone with selfish behaviors, regardless of the individual status in the UG (for enhancements of the advantageous status of the proposer, see e.g., Zak *et al.*, 2009, described above). For monetarily low but equal status, we previously found testosterone to decrease the willingness to forgo self-gains in order to increase the welfare of others (Ou et al., 2021; Wu et al., 2019). Even in a high and stable status, testosterone enhanced the willingness to compete for resources in order to avoid a status loss (Losecaat Vermeer et al., 2020). For a disadvantageous status, testosterone has been associated not only with increased rejection rates of offers that would create DI in the UG (Burnham, 2007; Zak et al., 2009) but also with enhanced competitive efforts for promoting social status (Josephs et al., 2006; Losecaat Vermeer et al., 2020; Mehta and Josephs, 2006). Here, we add to this literature by establishing a status enhancing role of testosterone in the absence of retaliation and strategic considerations.

Motivating our Hypothesis 2, recent evidence also demonstrated that testosterone promotes prosocial behaviors. For example, using a modified version of the UG, Dreher and colleagues found that men receiving testosterone provided more reward for partners who had repeatedly behaved generously to them beforehand (Dreher et al., 2016). Similar effects occurred also when males donating to a charity

organization were observed by others (Wu et al., 2020). In these studies, reputation and self-image play a central role either because of repeated interactions with the same person or because behavior is public. It is well known (Izuma, 2012; Nowak and Sigmund, 2005; Seinen and Schram, 2006) that reputation gains from prosocial behavior lead to resource (e.g., respect, network) gains in the long run, thereby enhancing social status. In our case, reputation gains were absent because all interactions were one-shot and private. Accordingly, being nice to the partner (either by counteracting advantageous inequality or enduring disadvantageous inequality) would provide little benefits for individuals pursuing higher social status. Thus, the relative effectiveness of selfish versus generous choices to enhance status may explain why testosterone facilitates one or the other behavior in different studies.

Several limitations of the current study should be noted. First, while the Fehr-Schmidt model was the best-fitting model and testosterone administration modulated inequality aversion, it is difficult to completely rule out the alternative account that testosterone also influenced the processing of self-reward, our design included a moderate but significant correlation between self-reward and payoff inequality. Hence, it would be beneficial for future studies to adopt a design in which self-reward and payoff inequality are decorrelated or orthogonal. Second, the current task may not perfectly reflect the complexities of real-world social behaviors, due to its simplification and lack of social interactions. Future studies may consider employing paradigms with more ecological social interactions. Third, only male participants were recruited in our study. However, in previous research, women receiving testosterone were more likely to propose offers with less advantageous inequality in the UG (Eisenegger et al., 2010). Thus, together with our findings, the effect of testosterone on advantageous inequality aversion might vary in men and women, which warrants further exploration. Furthermore, the current

study focused solely on measuring the choice behavior. To gain a deeper understanding of how testosterone causally affects fairness concern and to clarify its underlying psychological processes, it would be valuable to integrate other measures, such as self-reported motivation and measures of personality traits (e.g., selfishness, dominance).

In conclusion, our study contributes to the literature on the role of testosterone for social preferences by investigating both disadvantageous and advantageous inequality and by excluding confounds associated with the UG (e.g., retaliation and strategic). These findings shed new light on the function of testosterone in social decisions that involve a trade-off between personal interests and fairness concerns. Testosterone tuned aversion to advantageous and disadvantageous inequality, biasing men to prioritize their selfish interests, possibly in an attempt to enhance their social status in the future. The increased desire to be better off than others provides a novel specification of the association between testosterone and fairness concerns.

Acknowledgement

The authors are grateful to Dr. Jinting Liu and Ms. Yinhua Zhang for their help with the study.

Author contributions

YW, HY and PNT designed the study. JO and JL collected the data., JL, JO, and HY analyzed the data. JL and JO wrote the first version of the paper, YW, HY and PNT provided critical revisions. All authors approved the final version for submission.

Funding

Y.H. was funded by National Natural Science Foundation of China (32200853), Fundamental Research Funds for the Central Universities (2022ECNU-XWK-XK003), and Shanghai Natural Science Foundation (23ZR1418400). P.N.T. was supported by the Swiss NSF (100019_176016 and 100014_165884). Y.W. was supported by Start-up Fund for New Recruits (P0039779) and Research Institute for Sports Science and Technology (P0043556), The Hong Kong Polytechnic University. The funding sources had no further role in the study design, data collection, analysis, interpretation, and decision to submit this manuscript for publication.

Conflict of interests

We declare we have no competing interests.

References

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J.R., Brown, J.W., 2013. A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Decision, 1*(S), 8–23

Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting linear mixed-effects models using lme4. ArXiv Prepr. ArXiv14065823.

Boksem, M.A., 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. Psychol. Sci. 24, 2306–

2314.

Burnham, T.C., 2007. High-testosterone men reject low ultimatum game offers. Proc. R. Soc. B Biol. Sci. 274, 2327–2330.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A., 2017. Stan: A probabilistic programming language. J. Stat. Softw. 76.

Carré, J.M., Olmstead, N.A., 2015. Social neuroendocrinology of human aggression: examining the role of competition-induced testosterone dynamics. Neuroscience 286, 171–186.

Carré, J.M., Ortiz, T.L., Labine, B., Moreau, B.J., Viding, E., Neumann, C.S., Goldfarb, B., 2015. Digit ratio (2D: 4D) and psychopathic traits moderate the effect of exogenous testosterone on socio-cognitive processes in men. Psychoneuroendocrinology 62, 319–326.

Carré, J.M., Putnam, S.K., McCormick, C.M., 2009. Testosterone responses to competition predict future aggressive behaviour at a cost to reward in men. Psychoneuroendocrinology 34, 561–570.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Q. J. Econ. 117, 817–869.

Crockett, M.J., 2016. How formal models can illuminate mechanisms of moral judgment and decision making. Curr. Dir. Psychol. Sci. 25, 85–90.

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. Proc. Natl. Acad. Sci. 113, 11633–11638.

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.

Eisenegger, C., von Eckardstein, A., Fehr, E., von Eckardstein, S., 2013. Pharmacokinetics of testosterone and estradiol gel preparations in healthy young men. Psychoneuroendocrinology 38, 171–178.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191.

Fehr, E., Camerer, C.F., 2007. Social neuroeconomics: the neural circuitry of social preferences. Trends Cogn. Sci. 11, 419–427.

Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425, 785–791.

Fehr, E., Krajbich, I., 2014. Social preferences and the brain, in: Neuroeconomics. Elsevier, pp. 193–218.

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114, 817–868.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis.

Geniole, S.N., Proietti, V., Bird, B.M., Ortiz, T.L., Bonin, P.L., Goldfarb, B., Watson, N.V., Carré, J.M., 2019. Testosterone reduces the threat premium in competitive resource division. Proc. R. Soc. B 286, 20190720.

Gray, P.B., Straftis, A.A., Bird, B.M., McHale, T.S., Zilioli, S., 2020. Human reproductive behavior, life history, and the challenge hypothesis: a 30-year review, retrospective and future directions. Horm. Behav. 123, 104530.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. 3, 367–388.

Hermans, E.J., Ramsey, N.F., van Honk, J., 2008. Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. Biol. Psychiatry 63, 263–270.

Holper, L., Burke, C.J., Fausch, C., Seifritz, E., Tobler, P.N., 2018. Inequality signals in dorsolateral prefrontal cortex inform social preference models. Soc. Cogn. Affect. Neurosci. 13, 513–524.

Izuma, K., 2012. The social neuroscience of reputation. Neurosci. Res. 72, 283–288.

Josephs, R.A., Sellers, J.G., Newman, M.L., Mehta, P.H., 2006. The mismatch effect: when testosterone and status are at odds. J. Pers. Soc. Psychol. 90, 999.

Konovalov, A., Hu, J., Ruff, C.C., 2018. Neurocomputational approaches to social behavior. Curr. Opin. Psychol. 24, 41–47.

Kruschke, J., 2014. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. Fehr

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, 10–1016.

McAuliffe, K., Blake, P.R., Steinbeis, N., Warneken, F., 2017. The developmental foundations of human fairness. Nat. Hum. Behav. 1, 1–9.

Mehta, P.H., Josephs, R.A., 2006. Testosterone change after losing predicts the decision to compete again. Horm. Behav. 50, 684–692.

Morishima, Y., Schunk, D., Bruhin, A., Ruff, C.C., Fehr, E., 2012. Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. Neuron 75, 73–79.

Muller, M.N., 2017. Testosterone and reproductive effort in male primates. Horm. Behav. 91, 36–51.

Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437, 1291–1298.

Ou, J., Wu, Y., Hu, Y., Gao, X., Li, H., Tobler, P.N., 2021. Testosterone reduces generosity through cortical and subcortical mechanisms. Proc. Natl. Acad. Sci. 118, e2021745118.

Ruff, C.C., Ugazio, G., Fehr, E., 2013. Changing social norm compliance with noninvasive brain stimulation. Science 342, 482–484.

Seinen, I., Schram, A., 2006. Social status and group norms: Indirect reciprocity in a repeated helping experiment. Eur. Econ. Rev. 50, 581–602.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432.

Wagels, L., Radke, S., Goerlich, K.S., Habel, U., Votinov, M., 2017. Exogenous testosterone decreases men's personal distance in a social threat context. Horm. Behav. 90, 75–83.

Wagels, L., Votinov, M., Kellermann, T., Eisert, A., Beyer, C., Habel, U., 2018. Exogenous testosterone enhances the reactivity to social provocation in males. Front. Behav. Neurosci. 12, 37.

Wu, Y., Clark, L., Zilioli, S., Eisenegger, C., Gillan, C.M., Deng, H., Li, H., 2018. Single dose testosterone administration modulates emotional reactivity and counterfactual choice in healthy males. Psychoneuroendocrinology 90, 127–133.

Wu, Y., Zhang, Y., Ou, J., Hu, Y., Zilioli, S., 2020. Exogenous testosterone increases the audience effect in healthy males: evidence for the social status hypothesis. Proc. R. Soc. B 287, 20200976.

Wu, Y., Liao, J., Zilioli, S., Wu, Y., Deng, H., Li, H., Tobler, P.N., 2019. Testosterone administration increases social discounting in healthy males. Psychoneuroendocrinology 108, 127–134.

Zak, P.J., Kurzban, R., Ahmadi, S., Swerdloff, R.S., Park, J., Efremidze, L., Redwine, K., Morgan, K.,

Matzner, W., 2009. Testosterone administration decreases generosity in the ultimatum game. PloS One 4, e8330.

Zeileis, A., Wiel, M.A., Hornik, K., Hothorn, T., 2008. Implementing a class of permutation tests: the coin package. J. Stat. Softw. 28, 1–23.

Table 1 Model comparison for social preference models

Model	Form	LOOIC	*SE*
Fehr-Schmidt[a]	$\boldsymbol{U = M_{self} - \beta \cdot \max(M_{self} - M_{other}, 0) - \alpha \cdot \max(M_{other} - M_{self}, 0)}$	**4919.4**	**192.9**
General Inequality Aversion	$U = M_{self} - \alpha \cdot \lvert M_{self} - M_{other} \rvert$	8836.8	319.2
Aversion to Other's Gain	$U = M_{self} - \alpha \cdot M_{other}$	8244.4	344.2
Self-Other Tradeoff	$U = (1 - \alpha) \cdot M_{self} + \alpha \cdot M_{other}$	6230.7	312.3
Efficiency Concern	$U = (1 - \alpha) \cdot M_{self} + \alpha \cdot (M_{Self} + M_{other})$	5214.0	314.3

Note. The best-fitting model was highlighted in bold.

[a]We did not impose the constraints on the parameters as did in the original Fehr-Schmidt model, following the study by Holper and colleagues (2018).

LOOIC = Leave-One-Out Information Criterion. Lower LOOIC scores reflect better model fit. Bold font indicates the winning model. SE = standard error.