Testosterone and Cortisol Do Not Predict Rejecting Harm or Maximizing Outcomes in Sacrificial Moral Dilemmas: A Preregistered Analysis

Caleb J. Reynolds [a,b*]
Anastasia Makhanova [a,c]
Larissa Nikonova [a]
Lisa A. Eckel [a]
Paul Conway [a,d]

[a] Department of Psychology, Florida State University, 1107 W. Call St., Tallahassee, FL 32306, USA
[b] Department of Psychology, Wake Forest University, 1834 Wake Forest Rd., Winston-Salem, NC 27109, USA
[c] Department of Psychological Science, University of Arkansas, 216 Memorial Hall, Fayetteville, AR 72701, USA
[d] Department of Psychology, University of Portsmouth, King Henry I Street, King Henry Building, PO1 2DY, Portsmouth, United Kingdom

Accepted for publication, *Hormones and Behavior*, https://doi.org/10.1016/j.yhbeh.2021.105063

* Corresponding author. Email: reynolds.caleb.j@gmail.com

This work was supported by a Planning Grant (#042171) awarded to Paul Conway by the Florida State University Office of Research Development.

Abstract

Contemporary moral psychology explores the biological underpinnings of morality, including how neuromodulators influence moral judgment and decision making. Some studies suggest that higher circulating testosterone is associated with increased acceptance of sacrificial harm, such as killing one person to save five lives, consistent with utilitarian ethics and inconsistent with deontological ethics. However, most studies employ conventional analytic techniques that conflate concern about outcomes with reduced concern about sacrificial harm, many are statistically underpowered, and none examine potential regulating effects of cortisol. Therefore, we examined whether salivary concentrations of testosterone and cortisol jointly predict sacrificial dilemma judgments among a large sample of undergraduates (n = 199). We utilized an advanced cognitive modeling technique (process dissociation) to independently assess sensitivity to causing harm and maximizing outcomes, preregistering the prediction that higher testosterone would predict reduced harm-rejection rather than increased concern for outcomes, especially among people low in cortisol. However, neither testosterone, nor cortisol, nor their interaction predicted sacrificial dilemma response tendencies. Such findings raise questions about the robustness of past evidence suggesting links between testosterone and sacrificial dilemma judgments.

Keywords: testosterone, cortisol, moral judgment, dual-hormone hypothesis, process dissociation

Testosterone has important effects on human psychology, including modulating aggression and social interactions (e.g., Carré & Archer, 2018). Accordingly, researchers have examined links between testosterone and moral judgments of sacrificial dilemmas in which causing harm maximizes outcomes. Such work theorizes endogenous testosterone should predict willingness to cause sacrificial harm, i.e., to prefer utilitarian over deontological decisions (e.g., Lucas & Galinsky, 2015). Although some studies support this effect (e.g., Carney & Mason, 2010), the overall relationship appears mixed and complex (Brannon et al., 2019). Further complicating matters, studies on this topic often lack statistical power, employ conventional dilemma analyses treating harm rejection and outcome maximization as opposites, and ignore possible interactions between testosterone and cortisol (Mehta & Prasad, 2015). Therefore, we recruited a large sample to clarify the link between testosterone and sacrificial dilemma judgments, utilized process dissociation to disentangle harm rejection and outcome maximization responses, and tested for moderation by cortisol with preregistered analyses.

Sacrificial Dilemmas

Sacrificial dilemmas entail actions where causing limited harm can prevent greater suffering, such as killing a baby to save a village (Greene et al., 2001). Sacrificial actions align with *utilitarian* ethics that prioritize outcomes and violate *deontological* ethics that prioritize norms against harm, so researchers often describe accepting sacrificial harm as increased utilitarian or decreased deontological judgment. However, sacrificial decisions rarely reflect philosophical commitments, instead arising from a complex array of psychological processes, including concerns about rules, harm, and outcomes, as well as emotional concerns (or lack thereof) for others (e.g., Conway, in press).[1]

[1] Therefore, the terms *deontological* and *utilitarian* here denote neutral descriptors of decisions rather than philosophical commitments.

Accordingly, a growing body of work investigates the role of neurotransmitters and hormones in sacrificial judgments, especially those involved in processing concern for others (Crockett & Rini, 2015). In particular, some work suggests that testosterone might predict greater acceptance of sacrificial harm, suggesting reduced concern for victims or increased instrumental aggression (e.g., Carney & Mason, 2010). Consistent with this perspective, testosterone has been linked to decreased empathy, increased psychopathy, and stronger approach tendencies when facing threat—each of which predicts greater sacrificial harm (e.g., Lucas & Galinsky, 2015; Reynolds & Conway, 2018).

However, research directly examining testosterone and moral judgment is inconclusive. Some researchers report that higher endogenous testosterone is associated with greater endorsement of sacrificial harm (e.g., Carney & Mason, 2010; Chen et al., 2016; Montoya et al., 2013), but others report null relationships (e.g., Brannon et al., 2019), or marginal findings in the opposite direction (Arnocky et al., 2017). Accordingly, researchers have proposed several moderators, including sex (Armbruster et al., 2021; Gong et al., 2017), 2D:4D ratio (Montoya et al., 2013), hormonal birth control status (Armbruster et al., 2021), and dilemma features such as use of personal force (Carney & Mason, 2010) or the inevitability of harm (Chen et al., 2016; Montoya et al., 2013). Thus, although prior work suggests an association between testosterone and sacrificial harm—perhaps undergirded by reduced concern for victims—actual results are inconsistent.

Limitations of Past Work

We suggest past conclusions may be muddied by three issues: underpowered studies, measurement confounds, and potential moderating effects of cortisol. First, prior work on dilemmas and testosterone is often underpowered. The earliest and most-cited study to report an

association between testosterone and dilemma responses (Carney & Mason, 2010) observed effects equivalent to a correlation of $r \sim .2$. To detect this effect size with power of .80 (two-tailed test at $\alpha = .05$) requires 193 observations. Yet previous analyses examining endogenous testosterone and dilemmas average 65 participants (range: 19-181), few published studies approach this threshold, and key analyses often employ samples of 30 or fewer (cf. Brannon et al., 2019). Small samples decrease effect size estimate stability, increase false positive rates, and are more sensitive to researcher degrees of freedom (Button et al., 2013). Therefore, most studies measuring testosterone and moral judgments are underpowered, weakening confidence in positive findings and potentially generating the mixed findings.

Second, research on testosterone and dilemmas typically examines only dilemmas where causing harm maximizes outcomes (cf. Armbruster et al., 2021; Brannon et al., 2019). Such dilemmas allow for only conventional relative analyses treating harm avoidance and outcome maximization as opposites. Thus, such analyses cannot discern between insensitivity to harm and concern for outcomes, statistically suppressing measures related to both harm and outcome sensitivity (Reynolds & Conway, 2018). Instead, we employed Conway and Gawronski's (2013) dilemma battery which includes not only incongruent dilemmas where causing harm maximizes outcomes, but also congruent dilemmas where harm arguably does not. This battery allows for both a conventional analysis and a process dissociation (PD) analysis independently estimating the degree to which participant response patterns reflect harm-rejection (deontological) and outcome-maximization (utilitarian) tendencies. This can clarify whether testosterone influences dilemma judgments by increasing concerns about outcomes, reducing concerns about harm, or a more complex pattern. We expected testosterone to predict decreased sensitivity to harm (i.e., lower deontological parameter), rather than increased concern for outcomes (higher utilitarian

parameter). Such patterns would mirror the influence of gender, psychopathy, and empathic concern on sacrificial decisions assessed via PD (e.g., Reynolds & Conway, 2018).

Finally, prior dilemma research examined testosterone independent of other neuromodulators. However, complex social judgments likely reflect complex biological interactions. Recent research advances a *dual-hormone hypothesis*: testosterone predicts social behaviors at low levels of cortisol, but these relationships are attenuated by high cortisol (Mehta & Josephs, 2010; see Dekkers et al., 2019). Mehta and Prasad (2015) reviewed dual-hormone effects on status-relevant variables such as dominance leadership behaviors, antisociality, low empathic concern, and risk-taking—each of which predicts more 'utilitarian' judgments, typically via decreased sensitivity to outcomes (e.g., Conway & Gawronski, 2013; Lucas & Galinsky, 2015). The dual-hormone hypothesis suggests that unmeasured cortisol levels may partially explain previous inconsistent relationships between testosterone and sacrificial judgement. To test this possibility, we assessed both testosterone and cortisol.

The Current Work

We examined whether endogenous testosterone predicts moral dilemma responding in a large sample, using process dissociation to independently assess relationships between testosterone and outcome-maximizing and harm-avoidant response tendencies, and assaying both testosterone and cortisol to examine the dual-hormone hypothesis. We preregistered predictions that testosterone would predict lower harm-rejection—lower scores on the deontology parameter—but not increased sensitivity to outcomes (utilitarian parameter), and this effect would be attenuated by high endogenous cortisol.

Method

Including the Supplement, we report all measures, manipulations (none), and exclusions. All procedures were approved by the [UNIVERSITY REDACTED FOR REVIEW] Institutional Review Board, and appropriate APA ethical guidelines were followed. Materials, preregistration, data, and R analysis syntax are available at https://osf.io/e594m/?view_only=38c706331b84472cb9398ada85a7a37e. See Supplement (pp. 2-7) for further Method detail on the hormone assays, self-report measures, and process dissociation procedure.

Participants and Procedure

We recruited 249 undergraduates from a large public Southeastern US university for partial course credit between February 2018 and February 2019. The day before participation, we emailed participants to remind them of their appointment and remind them to not eat or drink (besides water) for an hour before their appointment, to not wear makeup or lip balm on their lips, and to not exercise prior to their appointment.[2] Participants came into the lab between 7:00am and 7:00pm, rinsed their mouths and provided saliva samples via passive drool before and after responding to questionnaires on a lab computer in a private room (only pretest samples were analyzed).

Prior to assaying hormones, we excluded two participants who did not complete informed consent and thus did not participate, 38 participants who failed an attention check, four who had saliva collection issues (e.g., inadequate sample or researcher collection error), and two who did not identify as either men or women, leaving 203 participants to assay. After assays, we excluded

[2] Most participants reported complying with these instructions: Noncompliance rates ranged from 2.5-7.5% for the three requests, and 175 of the final sample reported complying with all three requests. Patterns of results do not differ when only analyzing these 175 participants, so we retained the full 199 participants to maximize power. See Supplement (pp. 20-21) for these alternative analyses.

another three participants with extreme cortisol values (Z > 3 SDs on raw scores within the whole sample) and one participant with an extreme testosterone value (Z > 3 SDs on raw scores within gender). We analyzed data from 199 participants (116 women, M_{age}=19.52, SD=1.78), including 198 values for testosterone, 186 for cortisol,[3] and 185 for both. Participants were mostly White (74.9%), Hispanic (29.1%), Black (7.5%), Asian (2.5%), and Native American (1.5%).[4] To solidify confidence in our analyses, we preregistered the effective sample size (including exclusion criteria), hypotheses, and analyses for this study prior to any exploratory or confirmatory analyses (dated February 14, 2020).

Materials

We froze saliva samples at -20°C within an hour of collection. Before we assayed samples, they were thawed, centrifuged for 15min at 3000RPM, and the supernatant was refrozen in aliquots. We assessed testosterone and cortisol using commercial ELISA kits (Salimetrics, State College, PA). Consistent with known sex differences, Welch t-tests indicated that mean testosterone was higher for men (160.47pg/mL) than women (50.61pg/mL), t(103.02)=-17.59, p < .001, d=-2.68, but cortisol was not, (M_{men}=0.22µg/dL, M_{women}=0.19µg/dL), t(142.33)=-1.51, p=.134, d=-0.25. Accordingly, we sex-standardized testosterone but sample-standardized cortisol.

[3] As noted in the preregistration (https://osf.io/e594m/?view_only=38c706331b84472cb9398ada85a7a37e), we originally only assayed cortisol for 148 participants who started the study after 11:00am, reasoning to avoid including participants experiencing the diurnal surge in cortisol in the first hour after waking from sleep (Adam et al., 2017). However, a reviewer pointed out that this between-person variability is not inherently problematic, so we later assayed another ELISA kit's worth of participants (n = 38 randomly selected from the remaining participants with valid testosterone values who completed the study before 11:00am). The conclusions of this paper do not change based on whether we interpret the final or preregistered sample, so we retain the higher-powered final sample in the main text and present analyses with the preregistered sample size (n = 198 for testosterone, n = 148 for cortisol, n = 147 for both hormones) in the Supplement (p. 19).

[4] Participants could select multiple categories.

Participants completed the 20-dilemma PD battery (Conway & Gawronski, 2013), indicating whether each sacrificial action was *appropriate* or *inappropriate*. We calculated conventional 'utilitarian' judgments and the utilitarian (U) and deontology (D) parameters following Conway and Gawronski (see Supplement pp. 2-4 for more details). Finally, participants completed individual difference and demographic measures as criterion variables to compare dilemma results with past work: empathic concern, psychopathy, need for cognition and faith in intuition, moral identity, and aversion to performing and witnessing harm (Supplement, pp. 4-7). We present descriptive statistics and correlations for dilemma responses, hormone levels, and criterion variables in Table S1 in the Supplement (p. 13).

Results

Correlation and regression analyses suggested the dilemma battery performed as expected. We replicated previous relationships between the PD parameters and conventional judgments, gender, empathic concern, psychopathy, religiosity, and political conservatism, though need for cognition, moral identity, and harm aversion only partially replicated previous findings (Supplement, pp. 9-11). Likewise, testosterone and cortisol analyses replicated typical sex differences.

However, inconsistent with preregistered predictions, testosterone did not correlate with sacrificial dilemma decisions, whether assessed via conventional relative analyses, Pearson's r = -.08, p = .26, $CI_{95\%}$[-.22,.06], the U parameter, r = -.05, p = .52, $CI_{95\%}$[-.18,.09], or the D parameter, r = .06, p = .38, $CI_{95\%}$[-.08,.20]. Similarly, cortisol also did not significantly correlate with conventional relative judgments, r = -.04, p = .56, $CI_{95\%}$[-.19,.10], or the U, r = .01, p = .84, $CI_{95\%}$[-.13,.16] or D parameter, r = .07, p = .31, $CI_{95\%}$[-.07,.22]. These null patterns held when splitting the sample by gender (see Supplement p. 11).

Finally, we tested the preregistered dual-hormone prediction that high cortisol attenuates the relationship between testosterone and the deontological PD parameter by evaluating three linear models predicting conventional relative judgments and the U and D parameters from testosterone, cortisol, and their interaction (Table 1, Figure 1). Contrary to preregistered predictions, neither testosterone, cortisol, nor their interaction significantly predicted conventional judgments or either the U or D parameter, and sex did not moderate effects (Table 1). Moreover, to increase confidence in this result, we considered several exploratory alternative specifications of these models. We considered various more and less strict exclusions, used alternative ways of handling hormone outliers (e.g., Winsorizing, log-transformations), and controlled for effects of sex and time of day. We present results of these models in the Supplement (pp. 11-28). Results of these exploratory analyses were consistent with the conclusions from the preferred method reported here: We found no evidence that testosterone predicted dilemma responses, either independently or in conjunction with cortisol.

Discussion

Inconsistent with theory, preregistered hypotheses, and some past work, we found no significant relationship between endogenous testosterone and sacrificial dilemma judgments. This null effect emerged in a study employing a larger sample than previous work, employing both conventional and process dissociation analyses, and testing for moderation by cortisol. Such a null effect raises questions about interpreting previous findings. Past results are somewhat inconsistent, with some studies showing effects in the predicted direction and others failing to find effects. In particular, we highlight connections to two recent studies using advanced dilemma measurement techniques: studies done by Brannon and colleagues (2019) and Armbruster and colleagues (2021).

The current work replicates null findings from Brannon and colleagues (2019; *n* = 181 endogenous testosterone values), who also found no significant relationship between endogenous testosterone and sacrificial judgments on dilemmas where sacrificial action entails harm. Hence, the two highest-powered analyses of endogenous testosterone and dilemmas involving sacrificial action (Brannon et al. and the current work) both report null findings.[5]

However, our conclusions are at odds with recent work from Armbruster and colleagues (2021). These authors used a process dissociation approach to determine whether endogenous testosterone predicted harm-avoidant and outcome-maximizing response tendencies, but they separately analyzed men, women using hormonal contraceptives (HCs), and free cycling women. They reported that testosterone predicted higher U parameter scores (but not D parameter or conventional utilitarian judgment scores) among free cycling women (*n* = 44), predicted lower D parameter scores (though not U parameter or conventional scores) among men (*n* = 70), and did not predict moral dilemma responding among women using HCs (*n* = 32). We did not replicate any of these findings in our sample (see Supplement, pp. 21-22). Thus, even among studies employing similar advanced measurement techniques, there is disagreement about whether endogenous testosterone relates to sacrificial dilemma responding.

Therefore, to more definitively assess how hormones influence moral cognition, future work should focus on high-powered, preregistered tests, accounting for measurement issues and neuromodulator interactions raised here and elsewhere (for important design considerations in such studies, see Dekkers et al., 2019; Grebe et al., 2019; Knight et al., 2020). Future work may

[5] When additionally assessing dilemmas where action entails *preventing* sacrifice, Brannon and colleagues found that endogenous testosterone predicted reduced sensitivity to moral norms centered on concern for individuals. However, administering testosterone had the opposite effect of increasing sensitivity to such norms. Hence, the role of testosterone was not consistent between endogenous measures and exogenous administration. Moreover, considering actions that prevent sacrificial harm may entail a different psychology than considering actions that cause sacrificial harm, rendering inferences drawn from such dilemmas unclear for the current work (Conway, in press).

also consider assessing testosterone reactivity, rather than baseline testosterone (Zilioli & Bird, 2017).

Limitations and Future Directions

Null results are inherently difficult to interpret, as they could reflect a genuine null relationship between the variables of interest, or any other number of issues that may mask a true nonzero relationship. We consider here three possible interpretations of our null results, one related to statistical power, two related to the theoretical underpinnings of dual-hormone effects.

First, perhaps despite a larger than typical sample, our study was nonetheless underpowered. Although our tests of the simple effects of testosterone (n=199) were adequately powered to detect effects similar to prior studies (r = .2, Carney & Mason, 2010), our sample size for the interaction with cortisol (n = 185) provided only ~.66 power to detect $\Delta R^2 = .03$.[6] Nonetheless, our sample was larger than any published dilemma study directly measuring endogenous testosterone, and larger than 81% of dual-hormone studies (Dekkers et al., 2019; Grebe et al., 2019).[7] Thus, we suggest the current work still raises questions about the veracity of a link between testosterone and dilemma responding, either in conjunction with or regardless of cortisol.

Second, moral dilemma responding might be too distal of a variable on which to observe dual-hormone effects. Early formulations of dual-hormone effects focused on the potential regulatory effects of cortisol for links between testosterone and status-seeking behavior, such as acting in an (observer-rated) dominant or leader-like fashion (Edwards & Casto, 2013; Mehta & Josephs, 2010). Later research expanded to consider traits more distal to status-seeking

[6] ΔR^2=.03 approximates the largest meta-analytic estimate of a dual-hormone effect (Dekkers et al., 2019), making this power estimate a best-case scenario.
[7] Gong et al. (2017) used a larger sample (n = 654) but did not assess endogenous testosterone, instead inferring testosterone sensitivity from genotypes.

behaviors, such as empathy and psychopathy (e.g., Glenn et al., 2011; Zilioli et al., 2015; see Mehta & Prasad, 2015, for review)—traits also highly relevant to sacrificial moral judgments (e.g., Bartels & Pizarro, 2011; Gleichgerrcht & Young, 2013; Reynolds & Conway, 2018). Given this trait work, dual-hormone interpretations present a potential explanation for the inconsistent results in the testosterone and dilemma literature (see Brannon et al., 2019). Dilemma responses partly reflect social concerns, including conveying social information about the judge. Harm-acceptance signals competence, logical thinking, and leadership capability; harm-rejection signals warmth, morality, emotional feeling toward others, and suitability for close social roles (e.g., Bostyn & Roets, 2017a; Everett et al., 2016, 2018; Reynolds & Conway, 2021; Rom et al., 2017). Laypeople are aware of these perceptions and will strategically shift their public judgments to convey desired traits (Rom & Conway, 2018). Moreover, a perceived audience influences peoples' dilemma judgments (Bostyn & Roets, 2017b; Kundu & Cummins, 2013; Reynolds et al., 2019). Thus, it seems reasonable to consider potential testosterone effects on dilemmas in light of dual-hormone interpretations of status-seeking.

However, recent meta-analyses and reviews of dual-hormone effects question the generality of the dual-hormone hypothesis (Dekkers et al., 2019; Knight et al., 2020), arguing that testosterone × cortisol interactions relate more strongly to direct measures of status than to more indirect measures such as status-striving personality traits. Consequently, it may be that response tendencies on sacrificial dilemmas are simply too far removed from status-seeking behavior to consistently relate to testosterone and cortisol.

Finally, this latter possibility is amplified by potential social contextual constraints on dual-hormone effects. Links between testosterone and aggression or status seeking have long been discussed in relation to the challenge hypothesis (Archer, 2006; Carré & Archer, 2018),

which posits that when presented with certain social stimuli (e.g., social threat, status challenge, mating opportunity), testosterone increases will produce competitive and aggressive responses. The dual-hormone hypothesis was proposed to account for inconsistent support for this hypothesis, and indeed the original dual-hormone findings explicitly incorporated social status and social threat (Mehta & Josephs, 2010). Thus, dual-hormone effects might better be conceptualized as an effect on status-specific behaviors in specific social contexts, rather than broad associations across situations (Grebe et al., 2019). If so, we may have failed to a relationship between testosterone and cortisol and dilemma responses because participants' experiences were of a socially sparse setting: alone in a room by themselves, with no discernable status elements.[8]

If dual-hormone effects are contextually specific (rather than associated with general traits), participants in the present study might not have experienced the requisite social stimulation to exhibit a dual-hormone effect on dilemma judgments. One way to test this (perhaps in conjunction with testosterone reactivity) would be to put participants in a situation in which dilemma responses directly relate to status attainment. For example, Maner and Mead (2010) had participants complete self-report measures that would ostensibly be used to assign leadership for a subsequent task; a similar setup could be used with leadership ostensibly assigned based on dilemma responses. People associate endorsement of sacrificial harm with leadership and will selectively shift their judgments in this direction to convey competence (e.g., Rom & Conway, 2018). Perhaps in such a social context, participants higher on testosterone (perhaps in conjunction with low cortisol) might then endorse sacrificial harm more than those

[8] Indeed, we have elsewhere described the same setup used in this study as a "private" setting, absent social pressures or observations (Reynolds et al., 2019).

lower on testosterone. Thus, future research on testosterone and moral judgments should carefully consider the social contexts in which testosterone can affect downstream cognitions.

Conclusion

A large study with preregistered analyses failed to replicate previously reported relationships between endogenous testosterone and increased acceptance of sacrifice in moral dilemmas. This null effect emerged not only on conventional analyses but also process dissociation analyses, and there was no evidence for an interaction with cortisol as the dual-hormone hypothesis suggests. Such findings raise questions about the underlying effect. Future work should focus on high-powered, preregistered designs that consider the interrelatedness of hormone systems.

Table 1

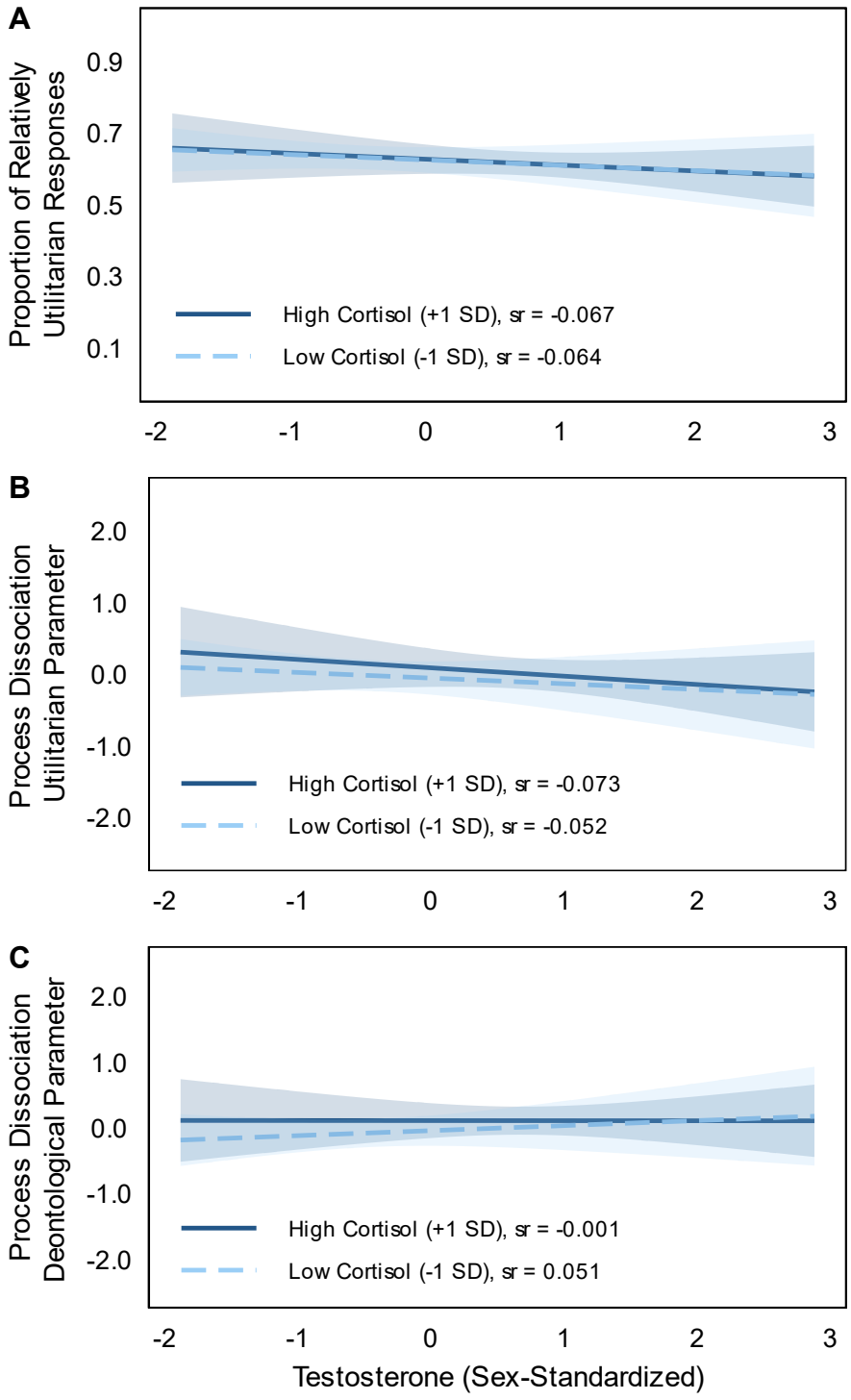
Regression Results for Models Predicting Conventional Utilitarian Judgments and the Process Dissociation U and D Parameters from Testosterone, Cortisol and Their Interaction

	Conventional Utilitarian Judgments			U Parameter			D Parameter		
Predictors	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**	**b**	**95% CI**	**p**
(Intercept)	**0.63**	**0.60 – 0.65**	**<0.001**	0.02	-0.15 – 0.18	0.845	0.02	-0.14 – 0.19	0.769
Testosterone	-0.02	-0.04 – 0.01	0.258	-0.10	-0.27 – 0.08	0.279	0.04	-0.14 – 0.21	0.672
Cortisol	0.00	-0.03 – 0.03	0.943	0.07	-0.12 – 0.26	0.449	0.08	-0.11 – 0.26	0.418
Testosterone × Cortisol	-0.00	-0.02 – 0.02	0.946	-0.02	-0.16 – 0.12	0.795	-0.04	-0.18 – 0.10	0.588
R^2 / R^2 adjusted		0.009 / -0.007			0.007 / -0.010			0.009 / -0.007	
(Intercept)	**0.51**	**0.44 – 0.59**	**<0.001**	-0.47	-0.97 – 0.03	0.065	**0.53**	**0.03 – 1.03**	**0.038**
Testosterone	-0.02	-0.10 – 0.06	0.559	-0.11	-0.64 – 0.42	0.693	0.06	-0.47 – 0.59	0.815
Cortisol	-0.07	-0.15 – 0.02	0.112	-0.32	-0.87 – 0.24	0.265	0.33	-0.23 – 0.89	0.243
Sex	**0.08**	**0.03 – 0.13**	**0.002**	**0.34**	**0.01 – 0.68**	**0.043**	**-0.35**	**-0.69 – -0.02**	**0.036**
Testosterone × Cortisol	0.00	-0.06 – 0.07	0.904	0.01	-0.42 – 0.43	0.977	-0.04	-0.47 – 0.38	0.838
Testosterone × Sex	0.00	-0.05 – 0.06	0.857	0.00	-0.36 – 0.36	0.993	-0.02	-0.38 – 0.35	0.933
Cortisol by Sex	0.05	-0.01 – 0.11	0.082	0.29	-0.10 – 0.68	0.141	-0.19	-0.58 – 0.19	0.330
Testosterone × Cortisol × Sex	-0.00	-0.05 – 0.04	0.892	-0.02	-0.30 – 0.27	0.915	0.00	-0.28 – 0.29	0.985
R^2 / R^2 adjusted		0.099 / 0.063			0.051 / 0.014			0.048 / 0.010	

Note. 185 observations. Statistically significant predictors ($\alpha = .05$)indicated in bold for emphasis. Testosterone is standardized within sex, cortisol within the whole sample.

Figure 1

Simple Slope Plots of the Relationships between Testosterone and Conventional Relative Judgments (A) *and the Standardized U and D Process Dissociation Parameters* (B & C), *Conditioned on Cortisol*



Note. The interaction was not significant ($\alpha = .05$) for any of the three outcome variables. *sr* denotes semi-partial correlation at the respective level of cortisol.

References

Adam, E. K., Quinn, M. E., Tavernier, R., McQuillan, M. T., Dahlke, K. A., & Gilbert, K. E. (2017). Diurnal cortisol slopes and mental and physical health outcomes: A systematic review and meta-analysis. *Psychoneuroendocrinology*, *83*, 25–41. https://doi.org/10.1016/j.psyneuen.2017.05.018

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*(3), 319–345. https://doi.org/10.1016/j.neubiorev.2004.12.007

Armbruster, D., Kirschbaum, C., & Strobel, A. (2021). Androgenic morality? Associations of sex, oral contraceptive use and basal testosterone levels with moral decision making. *Behavioural Brain Research*, *408*, 113196. https://doi.org/10.1016/j.bbr.2021.113196

Arnocky, S., Taylor, S. M., A. Olmstead, N., & Carré, J. M. (2017). The effects of exogenous testosterone on men's moral decision-making. *Adaptive Human Behavior and Physiology*, *3*(1), 1–13. https://doi.org/10.1007/s40750-016-0046-8

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154–161. https://doi.org/10.1016/j.cognition.2011.05.010

Bostyn, D. H., & Roets, A. (2017a). Trust, trolleys and social dilemmas: A replication study. *Journal of Experimental Psychology: General*, *146*(5), e1–e7. https://doi.org/10.1037/xge0000295

Bostyn, D. H., & Roets, A. (2017b). An asymmetric moral conformity effect: Subjects conform to deontological but not consequentialist majorities. *Social Psychological and Personality Science*, *8*(3), 323–330. https://doi.org/10.1177/1948550616671999

Brannon, S. M., Carr, S., Jin, E. S., Josephs, R. A., & Gawronski, B. (2019). Exogenous testosterone increases sensitivity to moral norms in moral dilemma judgements. *Nature Human Behaviour*, *3*(8), 856–866. https://doi.org/10.1038/s41562-019-0641-3

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Carney, D. R., & Mason, M. F. (2010). Decision making and testosterone: When the ends justify the means. *Journal of Experimental Social Psychology*, *46*(4), 668–671. https://doi.org/10.1016/j.jesp.2010.02.003

Carré, J. M., & Archer, J. (2018). Testosterone and human behavior: The role of individual and contextual variables. *Current Opinion in Psychology*, *19*, 149–153. https://doi.org/10.1016/j.copsyc.2017.03.021

Chen, C., Decety, J., Huang, P.-C., Chen, C.-Y., & Cheng, Y. (2016). Testosterone administration in females modulates moral judgment and patterns of brain activation and functional connectivity. *Human Brain Mapping*, *37*(10), 3417–3430. https://doi.org/10.1002/hbm.23249

Conway, P. (in press). Moral judgment. In *The Oxford handbook of social cognition* (2nd ed.). Oxford University Press.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, *104*(2), 216–235. https://doi.org/10.1037/a0031021

Crockett, M. J., & Rini, R. A. (2015). Neuromodulators and the (in)stability of moral cognition. In J. Decety & T. Wheatley (Eds.), *The moral brain: A multidisciplinary perspective* (pp. 221–235). The MIT Press.

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. https://doi.org/10.1016/j.neubiorev.2018.12.004

Edwards, D. A., & Casto, K. V. (2013). Women's intercollegiate athletic competition: Cortisol, testosterone, and the dual-hormone hypothesis as it relates to status among teammates. *Hormones and Behavior*, *64*(1), 153–160. https://doi.org/10.1016/j.yhbeh.2013.03.003

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216. https://doi.org/10.1016/j.jesp.2018.07.004

Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772–787. https://doi.org/10.1037/xge0000165

Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PLOS ONE*, *8*(4), e60418. https://doi.org/10.1371/journal.pone.0060418

Glenn, A. L., Raine, A., Schug, R. A., Gao, Y., & Granger, D. A. (2011). Increased testosterone-to-cortisol ratio in psychopathy. *Journal of Abnormal Psychology*, *120*(2), 389–399. https://doi.org/10.1037/a0021407

Gong, P., Fang, P., Yang, X., Ru, W., Wang, B., Gao, X., & Liu, J. (2017). The CAG polymorphism in androgen receptor (*AR*) gene impacts the moral permissibility of harmful behavior in females. *Psychoneuroendocrinology*, *80*, 74–79. https://doi.org/10.1016/j.psyneuen.2017.03.008

Grebe, N. M., Del Giudice, M., Emery Thompson, M., Nickels, N., Ponzi, D., Zilioli, S., Maestripieri, D., & Gangestad, S. W. (2019). Testosterone, cortisol, and status-striving personality features: A review and empirical evaluation of the Dual Hormone hypothesis. *Hormones and Behavior*, *109*, 25–37. https://doi.org/10.1016/j.yhbeh.2019.01.006

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Knight, E. L., Sarkar, A., Prasad, S., & Mehta, P. H. (2020). Beyond the challenge hypothesis: The emergence of the dual-hormone hypothesis and recommendations for future research. *Hormones and Behavior*, *123*, 104657. https://doi.org/10.1016/j.yhbeh.2019.104657

Kundu, P., & Cummins, D. D. (2013). Morality and conformity: The Asch paradigm applied to moral decisions. *Social Influence*, *8*(4), 268–279. https://doi.org/10.1080/15534510.2012.727767

Lucas, B. J., & Galinsky, A. D. (2015). Is utilitarianism risky? How the same antecedents and mechanism produce both utilitarian and risky choices. *Perspectives on Psychological Science*, *10*(4), 541–548. https://doi.org/10.1177/1745691615583130

Maner, J. K., & Mead, N. L. (2010). The essential tension between leadership and power: When leaders sacrifice group goals for the sake of self-interest. *Journal of Personality and Social Psychology, 99*(3), 482–497. https://doi.org/10.1037/a0018559

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. https://doi.org/10.1016/j.yhbeh.2010.08.020

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. https://doi.org/10.1016/j.cobeha.2015.04.008

Montoya, E. R., Terburg, D., Bos, P. A., Will, G.-J., Buskens, V., Raub, W., & van Honk, J. (2013). Testosterone administration modulates moral judgments depending on second-to-fourth digit ratio. *Psychoneuroendocrinology*, *38*(8), 1362–1369. https://doi.org/10.1016/j.psyneuen.2012.12.001

Reynolds, C. J., & Conway, P. (2018). Not just bad actions: Affective concern for bad outcomes contributes to moral condemnation of harm in moral dilemmas. *Emotion*, *18*(7), 1009–1023. https://doi.org/10.1037/emo0000413

Reynolds, C. J., & Conway, P. (2021). *"You will know them by their fruit": Expectations and perceptions of religious people's moral dilemma judgments* [Manuscript under review].

Reynolds, C. J., Knighten, K. R., & Conway, P. (2019). Mirror, mirror, on the wall, who is deontological? Completing moral dilemmas in front of mirrors increases deontological but not utilitarian response tendencies. *Cognition, 192,* 103993. https://doi.org/10.1016/j.cognition.2019.06.005

Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology*, *74*, 24–37. https://doi.org/10.1016/j.jesp.2017.08.003

Rom, S. C., Weiss, A., & Conway, P. (2017). Judging those who judge: Perceivers infer the roles of affect and cognition underpinning others' moral dilemma responses. *Journal of Experimental Social Psychology*, *69*, 44–58. https://doi.org/10.1016/j.jesp.2016.09.007

Zilioli, S., & Bird, B. M. (2017). Functional significance of men's testosterone reactivity to social stimuli. *Frontiers in Neuroendocrinology*, *47*, 1–18. https://doi.org/10.1016/j.yfrne.2017.06.002

Zilioli, S., Ponzi, D., Henry, A., & Maestripieri, D. (2015). Testosterone, cortisol and empathy: Evidence for the dual-hormone hypothesis. *Adaptive Human Behavior and Physiology*, *1*(4), 421–433. https://doi.org/10.1007/s40750-014-0017-x