Adaptive Human Behavior and Physiology
https://doi.org/10.1007/s40750-018-0098-z

ORIGINAL ARTICLE

Male Facial Attractiveness, Dominance, and Health and the Interaction between Cortisol and Testosterone

Tobias L. Kordsmeyer[1] · **Mareike Lohöfener**[1] · **Lars Penke**[1]

Received: 6 June 2018 / Revised: 10 July 2018 / Accepted: 11 July 2018

Abstract

Objectives The dual-hormone hypothesis suggests that associations of testosterone (T) with certain behavioral tendencies are stronger when cortisol (C) levels are low simultaneously. A range of studies provided supporting evidence for TxC interaction effects, for example on dominance and risk-taking behaviors. However, concerning perceptions of facial characteristics the evidence is mixed, with a recent study reporting a positive association between perceived facial dominance and T among men with higher C.

Methods We sought to further examine links of observer-rated facial attractiveness, dominance and health (based on photographs of $N = 165$ men) with baseline T, competition-induced T reactivity, and their interaction with baseline C.

Results There was some evidence that baseline T and the interaction of T reactivity with baseline C positively predicted facial dominance, however these were not robust when including control variables.

Conclusions Since no effects were found for perceived attractiveness and health, our results suggest that associations of perceived facial characteristics with baseline T, T reactivity and their interactions with baseline C are rather weak.

Keywords Testosterone · Cortisol · Faces · Attractiveness · Dominance · Health

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40750-018-0098-z) contains supplementary material, which is available to authorized users.

✉ Tobias L. Kordsmeyer
tob.kor@gmail.com

[1] Department of Psychology & Leibniz ScienceCampus Primate Cognition, University of Goettingen, Gosslerstr. 14, 37073 Goettingen, Germany

Introduction

From an evolutionary perspective, it is essential for humans to acquire mates and to reproduce (Puts 2016). To attract potential mates, men and women must signal their value as a sexual partner. One mechanism involved in acquiring mates is the endocrinological system. Hormones act as coordinators within the human body, affecting morphology, cognition and behavior (Roney 2016). For males, the steroid hormone testosterone (T) plays a key role in mating and intrasexual competition. High T levels have been associated with elevated mating effort (Archer 2006), offspring protection (van Anders et al. 2011) and development of secondary sexual characteristics such as muscle tissue and body hair (Vermeer et al. 2016). Men with high baseline T show more aggressive and dominant behavior (meta-analysis: $k = 42$ samples, overall $N = 9760$, Archer et al. 2005; $N = 69$, Sellers et al. 2007; 2 studies, overall $N = 151$, Mehta and Josephs 2010) and are rated as higher in extraversion and self-disclosure by observers ($N = 99$, Roney et al. 2007). Such behavioral tendencies are often inferred from an individual's facial characteristics, which are crucial cues in social interactions. For example, humans preferably interact with partners who have symmetric and average faces, assumingly because they are expected to be healthy and fit (for a review, see Little et al. 2011). It has been suggested that women also detect facial cues of men's hormonal concentrations and use these to assess mate attractiveness ($N = 39$ male stimuli, $N = 29$ female raters, Roney et al. 2006). Men's T levels are positively linked with women's ratings of facial attractiveness ($N = 74$ male stimuli, $N = 94$ female raters, Rantala et al. 2012; Roney et al. 2006), dominance (2 studies, $N = 50$ male stimuli, overall $N = 68$ male and female raters, Penton-Voak and Chen 2004; Roney et al. 2006), and may be related to health (as suggested by the immunocompetence handicap hypothesis, Folstad and Karter 1992; Foo et al. 2017).

Still, several studies found only weak associations of T with facial attractiveness and dominance. Penton-Voak and Chen (2004) used constructed and real male faces with high and low T levels ($N = 50$) and found no evidence for links between T and observer-rated attractiveness ($N = 40$ male and female raters), contrary to the above-cited studies. Similarly, Neave et al. (2003) found that among female ratings ($N = 36$) of facial photos of male college students ($N = 48$), T was not linked with attractiveness or dominance. Whereas Swaddle and Reierson (2002) found no direct link between T and observer-rated facial attractiveness, females did rate men with high T as higher in facial dominance ($N = 21$ male stimuli, $N = 50$ raters). One factor that might explain inconsistent and weak associations is that rather than baseline T, effects may be stronger for T reactivity, an acute increase in T levels, such as when engaging in a competition or mating opportunity (Archer 2006). Carré and Olmstead (2015) proposed that individual differences in behavior and traits may be more reliably related to T reactivity than to baseline T. Lefevre et al. (2013) investigated the association between male baseline T levels and T reactivity with facial width-to-height-ratio (fWHR, study 1: $N = 185$ men, study 2: $N = 92$ men), an objective measure supposedly mapping facial dominance. They found that fWHR was associated positively with T reactivity in the one study in which it was included, and baseline T in only one of the two studies. However, a later meta-analysis found these associations to be non-robust ($k = 9$ samples, $N = 1041$, Bird et al. 2016). If not objectively-measured facial traits, observer-perceived facial attractiveness and dominance may still be related to T reactivity. If these relationships turn

out to be robust, it would further support a role of acute T increases modulating behaviors and traits related to mate acquisition and intrasexual competition as suggested earlier (e.g., Archer 2006; Roney 2016).

A second explanation which has been proposed for weak links between traits and T is a buffering effect of the steroid hormone cortisol (C). Mehta and Josephs (2010) showed that across two studies (study 1: $N = 94$ men and women, study 2: $N = 57$ men), high baseline T was positively linked with observer-rated trait dominance only when baseline C levels were low simultaneously (in a leadership and competition context). For individuals with high C levels, the relationship between T and dominance was either non-significant or reversed (Mehta and Josephs 2010). Consequently, they proposed interacting effects of baseline T and baseline C (TxC interaction) on dominance-related behaviors and traits, termed the *dual-hormone hypothesis* (Mehta and Josephs 2010; for a review, see Mehta and Prasad 2015). Rantala et al. (2012) found that facial photos of men ($N = 74$) were rated as higher in attractiveness when their baseline T levels were high and C low simultaneously ($N = 94$ female raters). Facial dominance was also positively associated with baseline T only when C was low (Penton-Voak and Chen 2004). Conflicting evidence was provided by Kandrik et al. (2017), who investigated the baseline TxC interaction in association with perceptions of facial attractiveness, health and dominance. Seventy-three male and female participants rated five facial photos each of 45 heterosexual men (overall 225 facial photos). Whereas they found no effects of baseline T, baseline C or their interaction on health or attractiveness ratings, dominance ratings correlated with the TxC interaction. However, in contrast to the dual-hormone hypothesis and previous findings, men with high T were rated as higher in facial dominance only when simultaneously C levels were high. Thus, the exact direction and robustness of effects of TxC interactions on perceived facial dominance, attractiveness and health remain unclear.

Moreover, recent studies have found moderating influences of a prominent facial characteristic, beardedness, on perceptions of attractiveness and dominance in men. Beards have been suggested to be sexually selected, potentially being functional as threat displays in intrasexual competition (e.g., Puts et al. 2015). In several studies, effects of men's beardedness on perceptions of their facial attractiveness, dominance and health were shown (e.g., $N = 10$ men in four shaving conditions each, $N = 528$ male and female raters, Dixson and Brooks 2013; $N = 36$ men in four shaving conditions each, $N = 8520$ female raters, Dixson et al. 2016). Accordingly, in our study we examined our results' robustness by including beardedness as well as relationship status, age and BMI as control variables.

Aims and Hypotheses

In this study, we aimed to replicate and further investigate associations of men's baseline T, competition-induced T reactivity, also in interaction with baseline C, with observer-ratings of facial attractiveness, dominance and health. We hypothesized (1) high baseline T and (2) T reactivity to be positively linked with ratings of facial dominance, attractiveness and health. In line with the dual-hormone hypothesis, we hypothesized associations of (3a) baseline T and (4a) T reactivity with facial dominance, attractiveness and health to be attenuated by higher baseline C levels. Considering the findings of Kandrik et al. (2017), alternatively the faces of men with (3b) higher

baseline T and (4b) higher T reactivity may be perceived as more attractive, dominant and as healthier when C levels are simultaneously high.

Methods

Participants

One hundred sixty-five men participated in the study (mean age = 24.3 years, *SD* = 3.25, range = 18–34), of which 125 belonged to an experimental and 40 to a control group. From the participants, 90.4% were students (of which only two were enrolled as psychology students). On the 7-point Kinsey scale of sexual identity (1 = exclusively heterosexual to 7 = exclusively homosexual; Kinsey et al. 1948), the mean was 1.19 (*SD* = 0.46). None of the participants indicated to be taking any hormonal supplements or medication. All procedures received ethics approval from the Georg-Elias-Müller-Institute of Psychology's Ethics Committee (no. 111).

Procedure

Our study was divided into a pre- and a main session. The pre-session was led by a male experimenter. Participants provided informed consent and information on their relationship status and age. Anthropometric measures (body height using a stadiometer and weight using a digital scale, from which BMI was calculated), a first saliva sample (for baseline testosterone (T) and cortisol (C) measures) and standardized facial photographs were taken (and further measures not relevant to this study, for details see Kordsmeyer and Penke 2017). For the latter, participants stood in front of a white wall, directly faced the camera (Canon EOS 350D) with a distance of 2 m and were asked to show a neutral facial expression. Two photos were taken for each participant and the more suitable photo (regarding neutral expression and head angle) was chosen for the subsequent rating study. The main session was conducted a few days later and participants first provided a saliva sample for second baseline T and C measures and filled in questionnaires not relevant for this study (see Kordsmeyer and Penke 2017). In the experimental group, pairs of men engaged in a dyadic competition (four disciplines, mixture of cognitive and physical games, such as arm wrestling and table pinball soccer game, see Kordsmeyer and Penke 2017), while being supervised by an attractive female confederate. Both the competition and exposure to the attractive confederate were carried out to elicit a T response. The female confederate was told to interact naturally with the participants, while providing some verbal encouragement. Instead, in the control group, participants watched a neutral documentary video for the duration of the competition (ca. 12 mins, SFM Canada, 2013) and the procedure was led by a man, in order not to trigger T reactivity. Immediately after the competition or watching the documentary, and again 15–20 min later, participants provided a saliva samples to measure T reactivity, and filled in further questionnaires not relevant here (see Kordsmeyer and Penke 2017). Based on suggestions of a delay of 15–20 min for T responses to be detectable in saliva (Casto and Edwards 2016; Schultheiss et al. 2012), the two post-competition T measures can be interpreted as follows: The first can be seen as a measure of anticipatory reactivity (Marler et al. 2005) and a T increase during the competition's first minutes. The second post-sample can be interpreted as representing T reactivity during the full

competition phase and especially the two later disciplines (arm wrestling and turn-taking verbal fluency game). Both sessions were held between 2 pm and 6 pm to control for diurnal variation in hormonal levels (Idris et al. 2017; Stanton and Schultheiss 2009).

Hormonal Samples

For each sample, participants provided at least 2 ml of saliva via passive drool through a straw (Fiers et al. 2014; Schultheiss et al. 2012), which were immediately stored at −80 °C. Afterwards, saliva samples were shipped on dry ice to the Technical University of Dresden and analyzed for T and C levels using chemiluminescence immuno-assays with high sensitivity (IBL International, Hamburg, Germany). The intra- and inter-assay coefficients (CVs) are below 8% for C and below 11% for T. Outliers were winsorized to three *SD*s (n = 8 in the experimental, n = 1 in the control group, in accordance with Mehta et al. 2015). All participants were asked to not drink alcohol, exercise, take recreational or non-prescribed clinical drugs on the day of the study, to not ingest caffeine or sleep 3 h before, and to not smoke, brush their teeth or drink (except for water) 1 h before their appointment. Adherence to these instructions were assessed with a screening questionnaire (Stanton and Schultheiss 2009; for further details, see Kordsmeyer and Penke 2017). Because all T and C measures were positively skewed and violated the assumption of normality (Shapiro-Wilk test statistics <.94, ps < .001), all variables were log10-transformed (e.g., Mehta et al. 2015). For baseline T, we aggregated the two baseline measures to get a more reliable measure (e.g., Idris et al. 2017). T reactivity was calculated as difference scores, for the two post-competition samples separately (baseline T values from the main session subtracted from post-competition levels, Lobbestael et al. 2014; Roney et al. 2007). Finally, we excluded one participant from the experimental condition due to missing data for baseline T and C (decreasing sample size to overall N = 164, and n = 124 for T reactivity).

Rating Study

Facial photos of 164 target men (one photo was excluded due to issues with the photograph) were presented on computer screens in randomized order. Eleven male participants (age: M = 23.2 years, SD = 3.0, range 19–29) rated facial dominance (item: "How dominant is this man?") and health ("How healthy is this man?") on 11-point Likert-scales (dominance: −5 = "extremely submissive" to +5 = "extremely dominant", health: −5 = "extremely unhealthy" to +5 = "extremely healthy"). Twelve female participants (age: M = 25.2 years, SD = 7.1, range 19–44) rated facial short-term attractiveness (item: "How sexually attractive is this man?") and long-term attractiveness ("How attractive is this man for a long-term relationship?") on 11-point Likert-scales (for both: −5 = "extremely unattractive" to +5 = "extremely attractive"). The raters first previewed all facial photos for 0.5 s each. Additionally, for each target man raters were asked to what degree they knew him on a 3-point scale (1 = "not at all", 2 = "know him by sight", 3 = "well"). Data points where a rater indicated to know a given target man well (= 3) were excluded from subsequent analyses. Interrater agreements were high (Cronbach's $\alpha > 0.79$, see Table 1). Since mean ratings for short- and long-term attractiveness were highly correlated (r = .95), the average was used ("mean attractiveness"). To control for target men's beardedness, three trained, independent coders (two female, age range 18–29 years) judged the beardedness of all targets on a

Table 1 Descriptive statistics of all main variables

	M	*SD*	Cronbach's α
Age	24.3	3.3	
BMI	23.98	3.81	
Beardedness	2.59	2.29	.92
Baseline cortisol	0.45	0.21	
Baseline testosterone	1.89	0.15	
1st testosterone reactivity	0.05	0.12	
2nd testosterone reactivity	0.08	0.14	
Facial dominance	0.56	1.28	.79
Mean facial attractiveness	−1.35	1.43	.86/.87[a]
Facial health	0.51	1.39	.84

Note: $N = 124–164$; [a] short−/long-term attractiveness; Cronbach's α = interrater agreement (intercoder agreement for beardedness); baseline cortisol/testosterone: mean of two measures, log10-transformed and winsorized

11-point Likert-scale (0 =" clean-shaven", 10 = "longest, fullest beard", $\alpha = .92$), which were averaged to a beardedness variable.

Statistical Analyses

Linear models were used to investigate effects of T, C and their interaction on facial ratings, separately for baseline T (one model) and T reactivity (two models for the two reactivity measures). Predictors were T levels (baseline T or T reactivity), baseline C, and TxC interactions. Analyses for baseline T were run on the full sample ($N = 163$) with the mean of the two baseline measurements. Analysis for T reactivity were run on the experimental group only ($n = 124$), since only in this group the experimental condition was aimed to elicit a T response. In these models, the baseline C measure from the main session was used, because T reactivity measures were taken on this day. For robustness checks, we ran two separate linear models, firstly including target men's coded beardedness as a covariate, secondly including age, BMI and relationship status (1 = single/open relationship, 2 = committed relationship/engaged/married) as covariates, the latter because being in a romantic relationship has been found to attenuate T levels in men (e.g., van Anders and Watson 2006). Analyses were performed using R (R Core Team 2015).

Data Availability The data and analysis script associated with this research are available at osf.io/65tz4.

Results

Descriptive statistics for all main variables and bivariate correlations between facial ratings can be found in Tables 1 and S1, respectively. T reactivity was significantly different from 0 in the experimental condition for both post-competition samples, and significantly larger

compared to the control group for the first, but not the second, post-competition sample (Tables S5-S6).

For facial dominance, there was a significant positive main effect of baseline T ($\beta = 0.47$, $SE = 0.17$, $p = .048$, $\eta^2 = .04$; TxC interaction: $p = .10$; for detailed results, see Table 2), which was robust when controlling for beardedness, as well as when controlling for relationship status, age and BMI (Tables S2-S3). No further significant effects of baseline T, C or their interaction on perceived facial dominance, attractiveness or health were found (all unsigned βs < 1.65, ps > 0.10; Table 2).

In the experimental group, a negative main effect of T reactivity (for the 1st T reactivity sample only: $\beta = -0.34$, $SE = 0.14$, $p = .01$, $\eta^2 = .05$; 2nd: $\beta = 0.23$, $p = .11$; Tables 3) on facial dominance was detected. This negative main effect of T reactivity was attenuated by high baseline C levels (T reactivity x baseline C interaction, for both T reactivity samples, 1st: $\beta = 0.32$, $SE = 0.15$, $p = .03$, $\eta^2 = .04$; 2nd: $\beta = 0.36$, $SE = 0.14$, $p = .027$, $\eta^2 = .05$, Table 3; Fig. 1). These effects were robust in both control models, except for the effect of the T reactivity x baseline C interaction for the first T reactivity sample when controlling for age, BMI and relationship status ($p = .06$; Tables S4-S5). No further effects of T reactivity or its interaction with baseline C were found on men's perceived facial dominance, attractiveness or health in either of the T reactivity models (all unsigned ps > .06; Table 3). Effects of the control variables can be found in the supplementary (Tables S2-S5).

Discussion

The goal of this study was to investigate whether baseline testosterone (T) and competition-induced T reactivity, also in interaction with baseline cortisol (C), are

Table 2 Linear models of baseline TxC interaction, baseline T and baseline C predicting rated facial attractiveness, dominance and health

	β	SE	t	p	Partial η^2
Facial attractiveness					
Baseline T	0.19	0.18	1.10	.27	.01
Baseline C	−0.05	0.96	−0.05	.96	.00
Baseline TxC	−0.13	1.01	−0.13	.90	.00
Facial domininace					
Baseline T	0.47	0.17	2.72	.007	.04
Baseline C	1.41	0.95	1.48	.14	.01
Baseline TxC	−1.64	1.00	−1.64	.10	.02
Facial health					
Baseline T	0.10	0.18	0.56	.58	.00
Baseline C	0.56	0.99	0.56	.57	.00
Baseline TxC	−0.57	1.04	−0.55	.58	.00

Note: $N = 164$; T = testosterone; C = cortisol; for each aggregate measures from two samples; partial η^2 = partial eta-squared effect size. Model significance for facial attractiveness/dominance/health: $F_{3,159} = 2.49/3.94/0.13$, $p = .06/.01/.94$, adjusted $R^2 = .03/.05/-.02$

Table 3 Linear models of the T reactivity x baseline C interaction, T reactivity and baseline C predicting rated facial attractiveness, dominance and health (experimental group only)

	1st T reactivity					2nd T reactivity				
	β	*SE*	*t*	*p*	Partial η^2	β	*SE*	*t*	*p*	Partial η^2
Facial attractiveness										
T reactivity	−0.15	0.14	−1.11	.27	.01	−0.08	0.14	−0.54	.59	.00
Baseline C	−0.15	0.11	−1.37	.17	.02	−0.14	0.11	−1.24	.22	.01
T react. x baseline C	0.25	0.15	1.68	.10	.02	0.27	0.14	1.87	.06	.03
Facial dominance										
T reactivity	−0.34	0.14	−2.48	.01	.05	−0.23	0.14	−1.63	.11	.02
Baseline C	−0.16	0.11	−1.47	.14	.02	−0.17	0.11	−1.54	.13	.02
T react. x baseline C	0.32	0.15	2.16	.03	.04	0.36	0.14	2.51	.01	.05
Facial health										
T reactivity	−0.20	0.14	−1.49	.14	.02	−0.14	0.14	−0.97	.33	.01
Baseline C	0.03	0.11	0.24	.82	.00	0.02	0.12	0.15	.88	.00
T react. x baseline C	0.15	0.15	1.02	.31	.01	0.17	0.14	1.15	.25	.01

Note: $n = 124$; 1st/2nd T reactivity = testosterone reactivity from 1st/2nd post-sample; C = cortisol; react. = reactivity; partial η^2 = partial eta-squared effect size. Model significance for facial attractiveness/dominance/health, 1st T reactivity: $F_{3,119} = 1.04/2.13/1.15$, $p = .38/.10/.33$, adjusted $R^2 = .00/.03/.00$, 2nd T reactivity: $F_{3,119} = 1.82/2.16/0.85$, $p = .15/.10/.47$, adjusted $R^2 = .02/.03/.00$

associated with observer-ratings of men's facial attractiveness, dominance and health. The faces of men with higher baseline T were rated as more dominant, the effect of which was robust in two control models (including target men's beardedness as well as relationship status, age, and BMI). T reactivity was positively related to rated facial dominance when baseline C was high simultaneously (for both T reactivity measures), in line with a recently reported finding by Kandrik et al. (2017), but inconsistent with the dual-hormone hypothesis (e.g., Mehta and Josephs 2010). Still, the TxC interaction effect was at least partly robust in the two control models. No further effects of baseline

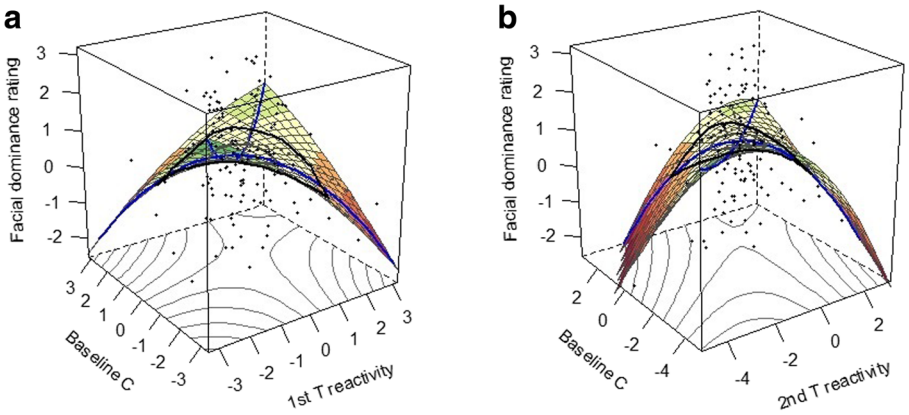

Fig. 1 Interaction between baseline cortisol and testosterone reactivity (panel A/B: first/second testosterone reactivity measure) on ratings of facial dominance

T, T reactivity, baseline C or any interactions on observer-rated attractiveness, dominance or health were detected.

The positive association between men's baseline T and observer-perceived facial dominance is in line with other studies (Penton-Voak and Chen 2004; Roney et al. 2006). Still, we could not replicate earlier findings on associations between baseline T and perceived facial attractiveness or health (e.g., Rantala et al. 2012; Roney et al. 2006; Penton-Voak and Chen 2004). Our null-findings for effects of baseline T on health and attractiveness ratings converge with some former studies (Kandrik et al. 2017; Neave et al. 2003; Penton-Voak and Chen 2004). Furthermore, besides an unexpected negative association between T reactivity and facial dominance we did not find effects of T reactivity on facial ratings, suggesting that at least for perceptions of facial characteristics associations may not be stronger for T reactivity than for baseline T (cf. Carré and Olmstead 2015).

Following the dual-hormone hypothesis, perceived facial characteristics such as dominance, attractiveness, or health may only be positively related to T when C levels are low (Mehta and Prasad 2015). In contrast, Kandrik et al. (2017) recently showed a different pattern. Ratings of facial dominance were positively related to baseline T among men with higher baseline C levels. In our study, we partly replicated their finding, in that the association of facial dominance with T reactivity was attenuated by low baseline C levels. However, we did not find an interacting effect for baseline T, leaving unresolved the question whether baseline T or T reactivity are more reliably associated with men's facial dominance, moderated by baseline C. In case either of these interacting effects can be replicated in subsequent studies and turn out to be robust, this hints at the possibility that C might not have a buffering but rather an enhancing effect on T. It has been suggested that C augments energy mobilization not only in threatening or stressful situations, but may also play a role in mating contexts, potentially facilitating mating effort (Jaremka and Collins 2017; van der Meij et al. 2010). Even though we did not find effects for facial attractiveness, the association for facial dominance, which is assumed to be functional in signaling to male rivals (Hill et al. 2017), may mean that the combination of high T reactivity and simultaneously high baseline C levels is implicated in intrasexual competitive and/or mating situations. Additionally, inconsistencies in findings on associations between hormonal levels, including TxC interactions, and perceptions of facial characteristics may partly be explicable by geographic variations in preferences (e.g., DeBruine et al. 2010), calling for further cross-cultural research. Moreover, since it has been stated that so far mechanisms of a TxC interaction can only be speculated about (e.g., Mehta and Prasad 2015), findings like ours and Kandrik et al.' (2017) stress the necessity of further investigating the exact neurophysiological origin of such an endocrinological interaction, and the exact directions of relationships between TxC and different traits and behaviors.

Generally, the many non-significant effects of baseline T, T reactivity and their interactions with baseline C raise the question to what extent hormones are related to signaling mate qualities and/or formidability via physical appearance and to what extent observers are able to detect hormonal cues from faces (Roney et al. 2006). Since the dual-hormone hypothesis was stated for behavioral measures and findings of associations between baseline T and behavior have been more consistent (Archer et al. 2005; Mehta and Josephs 2010; Mehta and Prasad 2015; Popma et al. 2007;

Roney et al. 2006; Sellers et al. 2007), it is possible that hormone levels are not as strongly linked with physical appearance as they are with behavior. Alternatively, some of our findings may represent false negatives, potentially explicable by overall small real effect sizes, which despite our comparatively large sample would benefit from even more statistical power. Therefore, it would be of interest to investigate hormonal effects in even larger samples, especially when considering the previously mixed and partly contrasting findings. Future research should also focus on associations between hormones and ratings of female faces; some first studies have already provided some promising results (Gonzalez-Santoyo et al. 2015; Rantala et al. 2013).

To conclude, we provide some further evidence for associations of hormone levels, and especially an interaction between testosterone and cortisol, with observer-perceptions of men's facial characteristics, but also provide further evidence of no associations for others. Relationships were found for facial dominance, but not attractiveness or health. These findings are in line with a recent study by Kandrik et al. (2017), in that testosterone positively predicted facial dominance when cortisol was high. Hence, they further question the previously reported direction of this interaction as suggested by the dual-hormone hypothesis. Our study paves exciting avenues for future research on hormonal associations, especially interactions between testosterone and cortisol, with not only facial, but also other traits and behaviors implicated in mating and competition.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews, 26*, 217–234. https://doi.org/10.1016/j.neubiorev.2004.12.007.

Archer, J., Graham-Kevan, N., & Davies, M. (2005). Testosterone and aggression: A reanalysis of book, Starzyk, and Quinsey's (2001) study. *Aggression and Violent Behavior, 10*, 241–261. https://doi.org/10.1016/j.avb.2004.01.001.

Bird, B. M., Cid Jofré, V. S., Geniole, S. N., Welker, K. M., Zilioli, S., Maestripieri, D., Arnocky, A., & Carré, J. M. (2016). Does the facial width-to-height ratio map onto variability in men's testosterone concentrations? *Evolution and Human Behavior, 37*, 392–398. https://doi.org/10.1016/j.evolhumbehav.2016.03.004.

Carré, J. M., & Olmstead, N. A. (2015). Social neuroendocrinology of human aggression: Examining the role of competition-induced testosterone dynamics. *Neuroscience Forefront Review, 286*, 171–186. https://doi.org/10.1016/j.neuroscience.2014.11.029.

Casto, K. V., & Edwards, D. A. (2016). Testosterone, cortisol and human competition. *Hormones and Behavior, 82*, 21–37. https://doi.org/10.1016/j.yhbeh.2016.04.004.

DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L., & Little, A. C. (2010). The health of a nation predicts their mate preferences: Cross-cultural variation in women's preferences for masculinized male faces. *Proceedings of the Royal Society of London B: Biological Sciences, 277*, 2405–2410. https://doi.org/10.1098/rspb.2009.2184.

Dixson, B. J., & Brooks, R. C. (2013). The role of facial hair in women's perceptions of men's attractiveness, health, masculinity and parenting abilities. *Evolution and Human Behavior, 34*, 236–241. https://doi.org/10.1016/j.evolhumbehav.2013.02.003.

Dixson, B. J. W., Sulikowski, D., Gouda-Vossos, A., Rantala, M. J., & Brooks, R. C. (2016). The masculinity paradox: Facial masculinity and beardedness interact to determine women's ratings of men's facial attractiveness. *Journal of Evolutionary Biology, 29*, 2311–2320. https://doi.org/10.1111/jeb.12958.

Fiers, T., Delanghe, J., T'Sjoen, G., Van Caenegem, E., Wierckx, K., & Kaufman, J. (2014). A critical evaluation of salivary testosterone as a method for the assessment of serum testosterone. *Steroids, 86*, 5–9. https://doi.org/10.1016/j.steroids.2014.04.013.

Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *The American Naturalist, 139*, 603–622. https://doi.org/10.1086/285346.

Foo, Y. Z., Nakagawa, S., Rhodes, G., & Simmons, L. W. (2017). The effects of sex hormones on immune function: A meta-analysis. *Biological Reviews, 92*, 551–571. https://doi.org/10.1111/brv.12243.

Gonzalez-Santoyo, I., Wheatley, J. R., Wellinga, L. L. M., Cárdenas, R. A., Jimenez-Trejo, F., Dawood, K., & Puts, D. A. (2015). The face of female dominance: Women with dominant faces have lower cortisol. *Hormones and Behavior, 71*, 16–21. https://doi.org/10.1016/j.yhbeh.2015.03.006.

Hill, A. K., Bailey, D. H., & Puts, D. A. (2017). Gorillas in our midst? Human sexual dimorphism and contest competition in men. In M. Tibayrenc & F. J. Ayala (Eds.), *On human nature: Biology, psychology, ethics, politics, and religion* (pp. 235–249). New York: Academic Press. https://doi.org/10.1016/B978-0-12-420190-3.00015-6.

Idris, F. P., Wan, Y., Zhang, X., & Punyadeera, C. (2017). Within day baseline variation in salivary biomarkers in healthy men. *A Journal of Integrative Biology, 21*, 74–80. https://doi.org/10.1089/omi.2016.0168.

Jaremka, L. M., & Collins, N. L. (2017). Cortisol increases in response to brief social exchanges with opposite sex partners. *Biological Psychology, 124*, 39–46. https://doi.org/10.1016/j.biopsycho.2017.01.001.

Kandrik, M., Hahn, A. C., Han, C., Wincenciak, J., Fisher, C. I., DeBruine, L. M., & Jones, B. C. (2017). Does the interaction between cortisol and testosterone predict Men's facial attractiveness? *Adaptive Human Behavior and Physiology, 3*, 275–281. https://doi.org/10.1007/s40750-017-0064-1.

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). Sexual behavior in the human male. Philadelphia: W. B. Saunders.

Kordsmeyer, T., & Penke, L. (2017, Dec. 19th). Effects of male testosterone and its interaction with cortisol on self- and observer-rated personality states in a competitive mating context. Retrieved from psyarxiv.com/wn5dc.

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior, 34*, 273–279. https://doi.org/10.1016/j.evolhumbehav.2013.03.005.

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society B, 366*, 1638–1659. https://doi.org/10.1098/rstb.2010.0404.

Lobbestael, J., Baumeister, R. F., Fiebig, T., & Eckel, L. A. (2014). The role of grandiose and vulnerable narcissism in self-reported and laboratory aggression and testosterone reactivity. *Personality and Individual Differences, 69*, 22–27. https://doi.org/10.1016/j.paid.2014.05.007.

Marler, C. A., Oyegbile, T. O., Plavicki, J., & Trainor, B. C. (2005). Response to Wingfield's commentary on "a continuing saga: The role of testosterone in aggression". *Hormones and Behavior, 48*, 256–258. https://doi.org/10.1016/j.yhbeh.2005.05.010.

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.

Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology, 56*, 88–99. https://doi.org/10.1016/j.psyneuen.2015.02.023.

Neave, N., Laing, S., Fink, B., & Manning, J. T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Family, 270*, 2167–2172. https://doi.org/10.1098/rspb.2003.2502.

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior, 25*, 229–241. https://doi.org/10.1016/j.evolhumbehav.2004.04.003.

Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., ... & Doreleijers, T. A. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological Psychiatry, 61,* 405–411. https://doi.org/10.1016/j.biopsych.2006.06.006.

Puts, D. (2016). Human sexual selection. *Current Opinion in Psychology, 7*, 28–32. https://doi.org/10.1016/j.copsyc.2015.07.011.

Puts, D. A., Bailey, D. H., & Reno, P. L. (2015). Contest competition in men. In D. M. Buss (Ed.), *The handbook of evolutionary psychology (Vol. 1: Foundations)* (pp. 385–402). Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781119125563.evpsych113.

Rantala, M. J., Moore, F. R., Skrinda, I., Krama, T., Kivleniece, I., Kecko, S., & Krams, I. (2012). Evidence for the stress-linked immunocompetence handicap hypothesis in humans. *Nature Communications, 3*, 694. https://doi.org/10.1038/ncomms1696.

Rantala, M. J., Coetzee, V., Moore, F. R., Skrinda, I., Kecko, S., Krama, T., Kivleniece, I., & Krams, I. (2013). Facial attractiveness is related to women's cortisol and body fat, but not with immune responsiveness. *Biology Letters, 9*, 1–4. https://doi.org/10.1098/rsbl.2013.0255.

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org. Accessed 8 June 2018.

Roney, J. R. (2016). Theoretical frameworks for human behavioral endocrinology. *Hormones and Behavior, 84*, 97–110. https://doi.org/10.1016/j.yhbeh.2016.06.004.

Roney, J. R., Hanson, K. N., Durante, K. M., & Maestripieri, D. (2006). Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society of London B: Biological Sciences, 273*, 2169–2175. https://doi.org/10.1098/rspb.2006.3569.

Roney, J. R., Lukaszewski, A. W., & Simmons, Z. L. (2007). Rapid endocrine responses of young men to social interactions with young women. *Hormones and Behavior, 52*, 326–333. https://doi.org/10.1016/j.yhbeh.2007.05.008.

Schultheiss, O. C., Schiepe, A., & Rawolle, M. (2012). Hormone assays. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *Handbook of research methods in psychology (Vol. 1: Foundations, planning, measures, and psychometrics)* (pp. 489–500). Washington DC: American Psychological Association.

Sellers, G. J., Mehl, M. R., & Josephs, R. A. (2007). Hormones and personality: Testosterone as a marker of individual differences. *Journal of Research in Personality, 41*, 126–138. https://doi.org/10.1016/j.jrp.2006.02.004.

Stanton, S. J., & Schultheiss, O. C. (2009). Hormonal correlates of implicit power motivation. *Journal of Research in Personality, 43*, 942–949. https://doi.org/10.1016/j.jrp.2009.04.001.

Swaddle, J. P., & Reierson, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proceedings of the Royal Society of London B: Biological Sciences, 269*, 285–2289. https://doi.org/10.1098/rspb.2002.2165.

van Anders, S. M., & Watson, N. V. (2006). Relationship status and testosterone in north American heterosexual and non-heterosexual men and women: Cross-sectional and longitudinal data. *Psychoneuroendocrinology, 31*, 715–723. https://doi.org/10.1016/j.psyneuen.2006.01.008.

van Anders, S. M., Goldey, K. L., & Kuo, P. X. (2011). The steroid/peptide theory of social bonds: Integrating testosterone and peptide responses for classifying social behavioral contexts. *International Society of Psychoneuroendocrinology, 36*, 1265–1275. https://doi.org/10.1016/j.psyneuen.2011.06.001.

van der Meij, L., Buunk, A. P., & Salvador, A. (2010). Contact with attractive women affects the release of cortisol in men. *Hormones and Behavior, 58*, 501–505. https://doi.org/10.1016/j.yhbeh.2010.04.009.

Vermeer, A. B., Riečanský, I., & Eisenegger, C. (2016). Competition, testosterone, and adult neurobehavioral plasticity. *Progress in Brain Research, 229*, 213–238. https://doi.org/10.1016/bs.pbr.2016.05.004.