



Do Testosterone and Cortisol Jointly Relate to Adolescent Dominance? A Pre-registered Multi-method Interrogation of the Dual-Hormone Hypothesis

Allison N. Shields¹ · Cassandra M. Brandes¹ · Kathleen W. Reardon² · Raul A. España¹ · Jennifer L. Tackett¹

Received: 14 December 2020 / Revised: 14 April 2021 / Accepted: 16 April 2021 /

Published online: 29 April 2021

© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Objective The goal of the present study was to extend the findings of the dual-hormone hypothesis (DHH) literature by assessing whether the interaction between testosterone (T) and cortisol (C) is associated with dominance in an adolescent sample via multiple methods of measuring T, C, and dominance, and with pre-registration of hypotheses and analyses.

Methods In a sample of 337 adolescents ($M_{age} = 14.98$, $SD = 1.51$; 191 girls) and their caregivers, hormonal assays were obtained from hair and saliva, and dominance behavior was assessed across four operationalizations (behavioral ratings in a leadership task, self- and caregiver reported dominance motivations, and self-reported social potency).

Results T and C main effects were generally null across hormone and dominance operationalizations, except that observer-rated dominance was negatively associated with salivary T, and social potency was positively associated with salivary T and negatively associated with salivary C. Support for the DHH was weak. Point estimates reflected a small negative $T \times C$ interaction for behavioral ratings of dominance, consistent with the DHH, whereas interaction effects for report-based dominance measures were close to zero or positive.

Conclusions The results contribute to a growing evidence base suggesting $T \times C$ interaction effects are variable across measures and methods used to assess hormones and dominance and highlight the need for comprehensive, multi-method examinations employing best practices in scientific openness and transparency to reduce uncertainty in estimates. Measurement of hormones and dominance outcomes vary across labs and studies, and the largely null results should be considered in that context.

Keywords Testosterone · Cortisol · Dual-hormone · Dominance · Adolescence

✉ Jennifer L. Tackett
jennifer.tackett@northwestern.edu

Extended author information available on the last page of the article

Aggression, risk-taking, and other types of disinhibitory behavior have been the focus of decades of empirical study due to their potential for deleterious outcomes. In particular, in seeking to explain why males exhibit higher levels of physical aggression and risk-taking than females, researchers have often focused on the role of androgens. Testosterone (T)—a steroid hormone product of the hypothalamic–pituitary–gonadal axis—has been highlighted as one potentially influential factor associated with status-relevant behaviors (e.g., aggression, dominance). Evidence for the T–aggression association appears to be robust in animal studies; however, evidence for the effect in humans is much less consistent (Archer, 2006; Carré & Archer, 2018; Wingfield et al., 1990). Mixed empirical findings suggest, at best, a weak positive correlation between T and status-relevant behaviors (Book et al., 2001; Rubinow & Schmidt, 1996). Indeed, a meta-analysis on the subject finds a significant but very small effect (Archer et al., 2005).

The dual-hormone hypothesis (DHH) has been offered as one way to account for inconsistent findings regarding the role of T on status-seeking behaviors in humans. As defined by Mehta and Josephs (2010), the DHH posits that T's association with dominance and aggression depends on levels of cortisol (C), a hormone product of the hypothalamic–pituitary–adrenal axis. Behaviorally, C is implicated in the stress response and associated with submissiveness, withdrawal, fear, and avoidance (Brown et al., 1996; Denson et al., 2009; Dickerson & Kemeny, 2004; Goldsmith & Lemery, 2000; Roelofs et al., 2009; Schulkin et al., 1998; Taylor et al., 2000), which are quite distinct from the behaviors associated with T (approach-related behaviors, antagonism, and risk-taking; Carré & McCormick, 2008; Mehta & Josephs, 2006; Oyebile & Marler, 2005; Platje et al., 2015; van Honk et al., 2004). Specifically, the DHH states that T and C jointly relate to socially dominant behavior such that higher levels of T are associated with more status-relevant behaviors, but only when C is low (Mehta & Josephs, 2010). When C is high, the DHH predicts that the effects of T would be inhibited or suppressed. Since it was first proposed, this model has been discussed extensively in the literature (see Knight et al., 2020; Mehta & Prasad, 2015; Sarkar et al., 2019), and the interaction between T and C has been linked to many status-relevant behavioral outcomes.

Proposed Mechanisms of the DHH

Support for the theoretical underpinnings of the DHH comes from a variety of research targeting potential mechanisms of the T×C interaction (Dekkers et al., 2019; Mehta & Josephs, 2010). As noted previously, behavioral manifestations of T and C are largely at odds with one another. These conflicting behavioral manifestations could potentially result in an interaction effect whereby high levels of one hormone override, or suppress, the other. Further support for the DHH can be drawn from evidence for T×C interactions at the molecular level – for example, high levels of C have been shown to actively suppress T expression at multiple molecular levels (Burnstein et al., 1995; Chen et al., 1997; Johnson et al., 1992; Tilbrook et al., 2000), and vice versa. This mechanistic process has been empirically observed at the cellular and genomic levels, as well (Viau, 2002). Thus, biological support for

T×C interaction has been drawn on to reinforce proposals for the DHH at behavioral levels.

Multiple studies have now found support for the DHH across a variety of behavioral outcomes, with fairly extensive recent coverage of these studies found elsewhere (Dekkers et al., 2019; Knight et al., 2020; Sarkar et al., 2019). Specifically, these studies typically find that associations between T and the outcome of interest are only observed when levels of C are low. Within this context, many status-relevant variables have been examined as the dependent variable, including social status and winning (Casto et al., 2019; Edwards & Casto, 2013; Ponzi et al., 2016; Sherman et al., 2016), social dominance (Mehta & Josephs, 2010; Pfattheicher, 2017), disinhibitory psychopathology and physical aggression (Dabbs et al., 1991; Geniole et al., 2011; Pfattheicher et al., 2014; Popma et al., 2007; Tackett et al., 2014), and risk-taking (Mehta et al., 2015; Ronay et al., 2018; Van den Bos, 2013).

Inconsistent Evidence Supporting the DHH

Despite this wide-ranging evidence supporting the DHH for a variety of status-relevant behaviors, other studies have shown reversed or null findings (e.g., Geniole et al., 2013; Grebe et al., 2019). Dekkers and colleagues (2019) recently conducted a meta-analysis examining the T×C interaction in predicting status, dominance, aggression, psychopathy, and risk-taking. This meta-analysis concluded that there was a significant T×C interaction in predicting these outcomes, although the effect size was very small ($r = -0.061$ overall; $r = -0.03$ for dominance). The authors failed to find sizeable moderation of this effect by participant sex, age, or population type. Further, Dekkers and colleagues (2019) found evidence of methodological weaknesses in the literature on the DHH, including low power, variability in operationalization of variables of interest, and analytic flexibility.

Potential Reasons for Inconsistency

A number of moderators not addressed by Dekkers et al. (2019) have been proposed to explain heterogeneity in the DHH literature. Prior studies have tested whether the DHH may only be detectable in targeted groups, while higher measurement error or restricted range of variance in other groups (e.g., lower T signal in females) may substantially decrease researchers' ability to detect dual-hormone effects. As early DHH studies included highly specific samples (e.g., male delinquents in Popma et al., 2007), some have proposed that sample stratification across different moderators (e.g., age, gender, clinical status) may better isolate T×C interaction effects, if and when they do exist, by increasing the ratio of signal to noise. Other moderators that have demonstrated this pattern include developmental considerations aside from chronological age (e.g., the influence of pubertal status on T×C coupling; Ruttle et al., 2015) or personality traits (e.g., disinhibitory personality traits; Tackett et al., 2014). However, many prior studies have lacked the robust sample sizes required to detect these complex interactions (Dekkers et al., 2019).

Inconsistencies in sample types and assays used to measure hormone levels may also explain heterogeneity in DHH findings. Most DHH studies have used immunoassay methods to measure hormone levels in single saliva samples. This may contribute to measurement error, as immunoassay methods may inflate very low T concentrations (Welker et al., 2016). Associations between immunoassays and liquid chromatography tandem mass spectrometry (LC–MS/MS) assays are weaker for T than for C, suggesting more precise assays should be used to measure T concentrations in particular. LC–MS/MS methods have a number of advantages over immunoassays, including higher specificity, greater matrix independence, and a wider analytical range (Field, 2013). Additionally, whereas single saliva samples are moderately stable when collected at the same time of day (Harden et al., 2016), single samples are subject to fluctuations due to mood or reactivity to current environmental states. Hair sampling may capture more trait-like individual differences in C and T levels, as hormones extracted from hair measure aggregate free hormone secretion over several months (Stadler et al., 2012). There is a need to examine possible DHH effects across these sample and assay types, as effects may be most trustworthy when using hair and LC–MS/MS assay methods. Other compelling explanations for non-replication of the DHH effect include poor construct validation and other measurement practices (e.g., Grahek et al., *in press*), low power, and analytic flexibility (e.g., Bakker et al., 2016). The DHH has been proposed to impact constructs ranging from leadership and dominance behaviors (Mehta & Josephs, 2010) to psychopathy (e.g., Welker et al., 2014), aggression (e.g., Geniole et al., 2011), and risk-taking (e.g., Mehta et al., 2015). While all of these constructs may overlap to some extent, they are not isomorphic, and construct confusion may result in overgeneralizing findings from one domain to another. Evidence also suggests that measurement practices in many areas of psychology are often poor, with studies using previously unvalidated measures (Flake et al., 2017; Grahek et al., *in press*) and selectively reporting measures or variables showing hypothesized effects (Simmons et al., 2011). While several recent studies that used pre-registered analysis plans have found evidence in support of the DHH (Kordsmeyer & Penke, 2019; Lausen et al., 2020; Ronay et al., 2018; Schild et al., 2020), additional sufficiently powered, multi-method, and pre-registered replications of this question may help to resolve discrepancies in this literature.

The Present Study

The goal of the present study was to evaluate the DHH literature broadly through a conceptual extension of many recent dual-hormone studies (see [Supplementary Table 1](#) for methodological differences between the present study and three other relevant dual-hormone studies). We aimed to assess whether the interaction between T and C is associated with dominance in an adolescent sample via multiple measures of T, C, and dominance, and with pre-registered hypotheses and analyses. The first of our three primary aims was to examine whether salivary T and C interact in predicting dominance in adolescents as measured by observer ratings of performance in a leadership task similar to that used in Mehta and Josephs (2010) Study 1. Our

second aim was to evaluate evidence for the DHH across different hormone samples (hair and saliva), and different operationalizations of dominance (self- and parent-reported dominance motivation and self-reported social potency). We hypothesized no main effects of either T or C on dominance, but we did not have specific predictions about whether or not we would observe significant $T \times C$ interactions. Our third aim was to evaluate evidence for a three-way $T \times C \times \text{gender}$ interaction to assess whether the $T \times C$ interaction differs between males and females, given mixed findings regarding gender differences in the broader literature. We hypothesized that there would not be a difference in the $T \times C$ interaction across genders. Our hypotheses and analytic decisions were pre-registered after data collection, but before data cleaning or analysis (<https://osf.io/5mzga>), and any deviations from this pre-registration are noted below in the *Deviations from the Pre-Registration* section.

Methods

Participants

The total sample consisted of 337 adolescents ($M_{age} = 14.98$ years, $SD = 1.51$ years, 57% female) and their caregivers (83.5% mothers). Adolescents were recruited via community outreach in a large metropolitan area in the midwestern United States. Flyers were distributed online, in local high schools, and in the community. Inclusion criteria were fluency in English for both adolescent and parent and an age range of 13–17 for adolescents (18-year-olds still in high school were also eligible). Exclusion criteria were the presence of an intellectual disability, neurodevelopmental disorder, or psychotic disorder in the adolescent. We excluded three participants who self-identified as transgender or gender nonbinary, as T values were winsorized and standardized within-sex (see *Data Analysis* section below). Parent-reported (or self-reported, if parent-reported data were unavailable) race/ethnicity of adolescents were as follows: 8.9% Asian, 10.7% Black, 6.2% Hispanic/Latinx, 62.0% White, 11.6% other/multiple races; 0.6% did not report their race/ethnicity.

Procedures

Data from the present study were collected as part of the Game Changers project, a broad study assessing many domains, including personality, psychopathology, and social functioning. Ethics approval was obtained by the Institutional Review Board at Northwestern University. Female participants who had reached menarche were scheduled to participate in the study during the first 10 days of their menstrual cycle, when hormone levels are most stable (Liening et al., 2010). All participants were instructed to refrain from eating and drinking for 2 h, as well as to refrain from smoking for 4 h prior to the lab visit. Participants were also instructed not to floss and not to eat dairy or anything containing live cultures the day of their visit.

Data from hair and saliva samples, dominance questionnaires, cognitive tasks, and pubertal development are included in the present study. A simplified timeline

and version of the full in-person lab protocol, including a list of all measures administered, can be found on the OSF page for this project (<https://osf.io/9n8gf/>). Participants reported to the lab and completed informed consent (parents) and assent (adolescents). Hair and saliva samples were collected from youth within the first 1.5 h of the visit, followed by cognitive tasks and computer-based questionnaires (including measures of dominance motivation, social potency, and pubertal status). Youth participants later completed a leadership task near the end of the 3.5-h visit.¹ Parental participation consisted of approximately 45 min of computer-based questionnaires. Youth and parents who did not finish questionnaires in the lab were sent a link to complete them online.

Hormone Samples

Standard collection, processing, and assaying procedures were used for salivary hormones (Schultheiss & Stanton, 2009). Passive drool saliva samples were collected between noon and sundown in order to account for diurnal variation in hormone levels (Kirschbaum & Hellhammer, 1994; Liening et al., 2010). After completing assent, adolescent participants were asked to rinse their mouths with water, then drink four ounces of water. Participants provided two baseline saliva samples: T1 was completed 30 min after their arrival to the lab and after completing sedentary questionnaires, and T2 was completed approximately 45 min thereafter. Participants passively drooled into a 2-mL IBL vial through a sanitary straw. Saliva samples were kept at room temperature until participants took a short break or until the study visit was complete (approximately 2 to 3 h) then frozen at -80 degrees Celsius. Samples were shipped in dry ice to the Technical University of Dresden where they were centrifuged for 5 min. Hormone concentrations were measured using chemiluminescence-immunoassays (IBL International, Hamburg, Germany). The intra- and inter-assay coefficients of variation were <9% and <12%, respectively, for both hormones.

Hair samples, 3 mm in diameter and at minimum 3 cm in length, were collected from participants' scalps, inferior to the cranial bones. Hair samples were stored at room temperature and were shipped to the Technical University of Dresden where T and C concentrations were assessed from the point of the hair closest to the scalp. Samples were washed with isopropanol and were incubated in methanol for 18 h. Samples were then centrifuged at 10,000 rpm for 2 min. Hormone concentrations

¹ Procedures for the present study included the Trier Social Stress Test (Kirschbaum et al., 1993), a task that reliably elicits an acute increase in C levels. Inclusion of this task is a theoretically-relevant potential moderator of dual-hormone effects, as acute C responses (rather than basal C) may moderate the association between T and status-seeking behavior in stress conditions (Prasad et al., 2017; Prasad, Knight, et al., 2019a, 2019b). However, in the present study, the stress-induction task occurred approximately 1.25–1.5 h before the leadership task, which allowed sufficient time for cortisol levels to return to baseline. This time also included a 15-min break during which participants enjoyed self-guided relaxation. C levels at the time of the final saliva sample ($M=7.68$, $SD=7.92$), which was collected approximately 15 min before the leadership task, were slightly lower than baseline (i.e., averaged T1 and T2) C levels ($M=9.01$, $SD=11.20$; $t(262)=3.47$, 95% CI [0.54, 1.95]). T levels at the time of the final saliva sample ($M=29.23$, $SD=38.12$) were not significantly different than baseline T levels ($M=31.91$, $SD=44.88$; $t(262)=1.39$, 95% CI [-1.11, 6.43]).

of these hair samples were measured using LC–MS/MS. Hair extracts with known hormone concentrations were included with every batch of LC–MS/MS analysis to ensure reliable and valid hormone quantification. Full wash and assay procedures are detailed in Gao et al. (2013).

“Leader–Follower” Leadership Task

Participants were taken into a room with the examiner and a confederate (“follower”) who was seated in front of blocks from the Wechsler Adult Intelligence Scale block design task (Wechsler, 2008).² A video camera facing the participant was put in plain view. Participants were explicitly informed that they would take part in a leadership task where they would be judged on their leadership ability in providing directions to the follower on building block designs. Participants stood directly behind the follower and were given the first design by the examiner. They were then instructed to prompt the follower to start a timer, after which they were to provide verbal directions on how the follower was to construct the blocks in front of them to match the picture on the card (which was only observable to the adolescent participant; the follower did not see the picture of the design). Once the block construction was complete, the participant prompted the follower to stop the timer, and the examiner recorded the time and the design for accuracy. Participants completed a total of four designs: two 2×2 designs (4 blocks) and two 3×3 designs (9 blocks).³

Measures

Descriptive statistics for all variables are listed in Table 1. Descriptive statistics by participant gender are in [Supplementary Table 2](#).

Observer Ratings of Adolescent Dominance

Participants’ dominance behavior during the leadership task was rated by the examiner and the confederate “follower.” The dominance scale (Mehta & Josephs, 2010, Study 1) consists of 19 items: engaged, bored (reversed), leader-like, energetic, confident, shy/timid (reversed), gave clear instructions, comfortable, assertive, directive, indecisive (reversed), dominant, comfortable giving instructions, nervous (reversed), stumbled over words (reversed), masculine, anxious (reversed), strong posture, and hesitant (reversed). Participants were rated on all items using a 7-point Likert scale ranging from (1) *Does not describe him/her at all* to (7) *Describes him/her very well*. Cronbach’s alpha for the observer-rated dominance scale was 0.96.

² Fifty-one different examiners and 61 different followers were used.

³ While our leadership task was modeled after that used in Mehta and Josephs (2010) Study 1, there were important differences, including age and gender differences in leader and follower in the present study and fewer designs completed. Further, while participants were explicitly told they would be judged on leadership ability, they were not explicitly told their leadership skills would be compared to other participants.

Table 1 Descriptive Statistics

Variable	N	Mean	SD	Range	Skew	Kurtosis
Salivary Cortisol Time 1 (nmol/l)	271	9.75	7.37	1.19 – 44.03	2.18	5.73
Salivary Cortisol Time 2 (nmol/l)	270	7.25	4.76	1.21 – 30.57	1.83	4.48
Salivary Testosterone Time 1 (pg/ml)	271	32.08	37.23	0.08 – 195.69	1.74	2.94
Salivary Testosterone Time 2 (pg/ml)	270	27.48	33.91	0.00 – 177.36	1.98	3.88
Hair Cortisol (pg/mg)	219	15.70	34.44	0.10 – 228.00	4.54	22.32
Hair Testosterone (pg/mg)	219	0.76	1.33	0.10 – 9.34	4.52	23.97
AMS Dominance—Youth report	309	0.67	0.27	0.00 – 1.00	-0.64	-0.38
AMS Dominance—Parent report	272	0.66	0.29	0.00 – 1.00	-0.53	-0.76
MPQ-BF Social Potency	238	2.60	0.43	1.44 – 3.61	-0.06	-0.14
Observer-Rated Dominance	314	3.78	0.91	1.53 – 5.61	-0.20	-0.86
Leader–Follower Design 1 (seconds)	331	76.00	38.11	22.28 – 180.00	1.32	1.16
Leader–Follower Design 2 (seconds)	328	84.02	39.23	21.94 – 180.00	0.94	0.09
Leader–Follower Design 3 (seconds)	330	133.04	50.20	44.19 – 240.00	0.62	-0.47
Leader–Follower Design 4 (seconds)	327	143.63	55.51	41.82 – 240.00	0.33	-0.93
Leader–Follower Averaged Z-Score	331	0.03	0.79	-1.55 – 2.33	0.54	-0.25
Digit Span Forward	336	10.18	2.09	0 – 16	-0.55	2.59
Digit Span Backward	336	6.27	2.25	0 – 13	0.20	0.14
Trail-Making Task Part B (seconds)	328	70.24	31.40	19.69 – 246.36	2.11	7.18

Cortisol and testosterone values presented represent winsorized values

Consistent with diurnal rhythms, salivary cortisol and testosterone decreased significantly from Time 1 to Time 2 (both $ps < .001$)

SD Standard Deviation, AMS Achievement Motivation Scale, MPQ-BF Multidimensional Personality Questionnaire – Brief Form

Dominance Motivation

Dominance motivation was measured using both self- and informant-report versions of the Achievement Motivation Scale (AMS; Cassidy & Lynn, 1989). The AMS is a 49-item scale designed to measure different motivations for achievement. Participants and parents responded dichotomously (*Yes* or *No*) on the 7-item Dominance scale (e.g., *I think I would enjoy having authority over people* or *I think they would enjoy having authority over other people*). Cronbach's alpha for the Dominance scale was 0.71 for youth report and 0.77 for parent report.

Social Potency

Social potency was measured using the 18-item Social Potency subscale of the Multidimensional Personality Questionnaire Brief Form (MPQ-BF; Patrick et al., 2002; Tellegen, 1982). The MPQ-BF is a 155-item questionnaire designed to measure personality from a psychobiological framework. Adolescent participants responded to the Social Potency scale (e.g., *I am quite good at convincing others to see things my*

way) on a 4-point Likert scale ranging from (1) *Definitely true* to (4) *Definitely false*. Cronbach's alpha for the youth-report Social Potency subscale was 0.81.

Covariates

Cognitive Performance. To maintain consistency with Mehta and Josephs (2010) Study 1, cognitive performance was used as a covariate in regression analyses, assessed via the following metrics.

Leadership Task Time. When the outcome variable was observer-rated dominance, cognitive performance was operationalized as the amount of time, in seconds, it took for participants to complete the leadership task. Each participant completed 4 designs, and the third and fourth designs were expected to take longer to complete than the first and second. We created z -scores for a participant's time on each design; the participant's average z -score across the four designs was used as the time-based cognitive performance score.

Executive Function Tasks. Because our data collection protocol also included decontextualized tasks indexing cognitive performance, we used scores on executive function tasks as cognitive performance covariates for survey-based outcome variables. We also used these tasks as cognitive performance covariates for the observer-rated dominance outcome variable as a robustness check. Executive function tasks assessed working memory maintenance (Digit Span Forward; Lahey et al., 2004), working memory updating (Digit Span Backward; Lahey et al., 2004), and set-shifting (Trail Making Test Part B; Reitan et al., 1992). The Digit Span Forward and Digit Span Backward scores were computed by summing the number of correctly recalled digit sequences. Performance on the Trail Making Task was operationalized as the amount of time, in seconds, it took for participants to complete Part B.

Testing Time of Day. Salivary cortisol levels vary diurnally (Kirschbaum & Hellhammer, 1994; Liening et al., 2010), so time of saliva sample collection (i.e., the midpoint between T1 and T2 collection times) was used as a covariate in regression analyses using salivary hormones, consistent with Mehta and Josephs (2010) Study 1. However, others suggest that cortisol assessment should be anchored on waking time, rather than sample collection time, as diurnal rhythms are more affected by sleep–wake cycles than by light–dark cycles (Adam & Kumari, 2009). In the present study, there were small negative correlations between time since waking (computed as the number of minutes between waking and sample collection) and salivary C (T1 $r = -0.02$, 95% CI [-0.14., 0.10]; T2 $r = -0.04$, 95% CI [-0.16, 0.08]) and T (T1 $r = -0.06$, 95% CI [-0.18, 0.06]; T2 $r = -0.07$, 95% CI [-0.19, 0.05]). Thus, we conducted additional robustness analyses using wake time as a covariate rather than sample collection time.

Pubertal Development. Pubertal development was rated using the Pubertal Development Scale (PDS; Petersen et al., 1988), a self-report measure designed to assess pubertal status in male and female adolescents. Males and females rated their growth in height, appearance of body hair, and changes in their skin using a 4-point scale, ranging from (1) *Has not yet started* to (4) *Seems completed*. Male participants also rated their growth of facial hair and their voice deepening, while female participants rated their breast development and the onset of menstruation. Onset of menstruation was measured using the responses *Yes* or *No*. Dimensional composite PDS scores were used in analyses ($M_{female}=3.16$, $M_{male}=2.81$); most female participants ($n=158$) were in the advanced pubertal or post-pubertal stages, whereas most male participants ($n=119$) were in the mid- or advanced pubertal stages of development.

Data Analysis

Some participants were missing either saliva ($n=66$) or hair ($n=118$) samples due to refusal or insufficient material (i.e., very short hair) for assays. The present sample reflects all participants who had either hair ($n=219$) or salivary ($n=271$) hormone data (total $N=337$). The number of participants with data for each variable used in the present study can be found in Table 1.

The lower sensitivity limits for salivary C and T are 0.11 nmol/L and 1.8 pg/ml, respectively. In the present study, no samples fell below the limit for salivary C, whereas 9 samples fell below the limit for salivary T. Excluding participants whose salivary T values were below the lower limit of detectability did not substantially affect results (see [Supplementary Table S6](#)), so we retained the salivary T levels provided for these 9 samples. The lower sensitivity limit for hair T and C is 0.1 pg/mg; samples below this threshold were given values of 0.1 pg/mg for hair T ($n=45$) or C ($n=2$).

For both salivary and hair hormones, we winsorized extreme C and T values (i.e., replaced with the highest value within 3 standard deviations of the sample mean). For T only, we winsorized values separately for males and females to account for sex differences in T levels. After winsorization, we averaged salivary T and C values obtained at T1 and T2. Pearson r s were 0.89 and 0.69 for associations between T1 and T2 salivary C and T, respectively. As expected, T and C values were skewed, so we used log-transformed values in analyses. Analyses were also performed with non-log-transformed T and C values as a robustness check.

Hierarchical linear regression models were estimated to examine primary study aims, including main effects and interaction effects. Eight hierarchical regression models were estimated to examine evidence for the DHH using a) salivary and b) hair hormones (T and C) as predictors, predicting dominance as measured by a) performance on the leadership task, b) parent-reported AMS Dominance, c) self-reported AMS Dominance, and d) self-reported MPQ Social Potency. Dominance was entered as the dependent variable for all regression models. Covariates (time of day for salivary hormone analyses and cognitive performance for all analyses) were entered in Step 1. Gender, T, and C were entered in Step 2; the T×C interaction term was entered in Step 3; finally, the T×gender, C×gender, and T×C × gender interaction terms were entered in Step 4.

Results

Deviations from the Pre-Registration

Analyses were conducted as pre-registered (<https://osf.io/9n8gf/>), with a few exceptions. First, we conducted a number of validity checks on our hormone data (e.g., examining whether T values were associated with gender, age, and pubertal status). Second, in an effort to more fully clarify associations between hormone outcomes and relevant covariates (age, gender, and pubertal status), we conducted a set of preliminary analyses in which hormonal outcomes were regressed on these covariates. Third, because hormonal concentrations change dramatically across pubertal development (Shirtcliff et al., 2009), we conducted an additional robustness check by controlling for pubertal status in regression models. Fourth, we conducted analyses excluding a subset of participants whose salivary T values were below the lower limit of detectability. Fifth, to maintain consistency with other DHH studies, we analyzed regression models separately for males and females. Output for all non-pre-registered analyses can be found on the OSF page for this project (<https://osf.io/9n8gf/>) and is explicitly labeled as such.

Preliminary Analyses

Hormonal outcomes were first regressed on relevant covariates: Age, Gender, Gender \times Age, and pubertal development (Supplementary Table 3). In initial models, Age² and Age² \times Gender effects were estimated as nonsignificant for all outcomes and were subsequently dropped to produce more interpretable main effects. Similarly, the Gender \times Age effect was estimated as nonsignificant for both hair outcomes and for salivary C and was dropped from those models. Results indicated that salivary C and both salivary and hair T concentrations were significantly higher in male participants, whereas hair C concentrations did not differ by gender. For salivary T, the gender difference significantly increased with age. Salivary T levels were also positively associated with age and pubertal development, whereas salivary C and hair T and C levels were not significantly associated with age or pubertal development. Given robust gender differences in testosterone levels, we standardized hair and salivary T within sex so that high levels indicate high T relative to other same-sex adolescents.

Correlations between all variables are presented in Supplementary Table 4a and are presented separately by participant gender in Supplementary Table 4b. Consistent with previous research (Mehta et al., 2008; Mehta & Josephs, 2010, Study 1), salivary T and C levels were positively correlated ($r=0.34$, $p<0.001$). Hair T and C levels were also positively correlated ($r=0.28$, $p<0.001$). Salivary and hair C were uncorrelated ($r=-0.03$, $p=0.68$), whereas salivary and hair T were negatively correlated ($r=-0.18$, $p=0.02$). Dominance variables were positively correlated, with associations ranging from $r=0.17$ ($p=0.01$) between observer-rated dominance and parent-reported AMS Dominance to $r=0.62$ ($p<0.001$) between youth-reported AMS Dominance and youth-reported MPQ Social Potency.

Dual-Hormone Hypothesis

As pre-registered, we conducted eight hierarchical regression analyses to evaluate evidence for the DHH. A negative $T \times C$ interaction term implies that T is positively associated with dominance only when C is low, supporting the DHH. Here, we present only T and C main effects, $T \times C$ interaction effects, and $T \times C \times$ gender interaction effects. Model results can be found in Table 2, and models run separately by participant gender can be found in Supplementary Table 5.

Observer-Rated Dominance

For salivary hormones, there was a small main effect of T on observer-rated dominance, such that adolescents with lower levels of T were rated as more dominant, on average ($\beta = -0.13$, 95% CI [-0.23, -0.03], $p = 0.02$). Main effects of C ($\beta = 0.07$, 95% CI [-0.10, 0.24], $p = 0.24$) and gender ($\beta < 0.01$, 95% CI [-0.19, 0.19], $p > 0.99$) on observer-rated dominance were nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta = -0.12$, 95% CI [-0.28, 0.05], $p = 0.55$, $\Delta R^2 < 0.01$; Fig. 1a). The $T \times C \times$ gender interaction was also nonsignificant ($\beta = 0.52$, 95% CI [0.19, 0.85], $p = 0.07$, $\Delta R^2 = 0.02$). However, in models run separately by participant gender, the direction of the $T \times C$ interaction was negative in males ($\beta = -0.42$, 95% CI [-0.70, -0.14], $p = 0.17$) and positive in females ($\beta = 0.28$, 95% CI [0.08, 0.49], $p = 0.27$).

For hair hormones, main effects of T ($\beta = 0.07$, 95% CI [-0.04, 0.18], $p = 0.27$), C ($\beta = -0.03$, 95% CI [-0.11, 0.06], $p = 0.65$), and gender ($\beta = 0.02$, 95% CI [-0.22, 0.26], $p = 0.74$) on observer-rated dominance were nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta = -0.15$, 95% CI [-0.23, -0.07], $p = 0.14$, $\Delta R^2 < 0.01$; Fig. 2a). The $T \times C \times$ gender interaction was also nonsignificant ($\beta = -0.27$, 95% CI [-0.46, -0.08], $p = 0.20$, $\Delta R^2 = 0.01$), though the $T \times C$ interaction was positive in males ($\beta = 0.07$, 95% CI [-0.11, 0.25], $p = 0.69$) and negative in females ($\beta = -0.22$, 95% CI [-0.32, -0.12], $p = 0.10$).

Parent-Reported AMS Dominance

For salivary hormones, main effects of T ($\beta = -0.05$, 95% CI [-0.09, -0.005], $p = 0.54$), C ($\beta = 0.02$, 95% CI [-0.05, 0.09], $p = 0.79$), and gender ($\beta = 0.06$, 95% CI [-0.02, 0.14], $p = 0.38$) on parent-reported dominance were nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta = 0.03$, 95% CI [-0.03, 0.10], $p = 0.89$, $\Delta R^2 < 0.01$; Fig. 1b) and the direction of this effect was equivalent in males and females. The $T \times C \times$ gender interaction was also nonsignificant ($\beta = 0.04$, 95% CI [-0.09, 0.18], $p = 0.90$, $\Delta R^2 < 0.01$).

For hair hormones, main effects of T ($\beta = 0.03$, 95% CI [-0.01, 0.08], $p = 0.66$), C ($\beta = 0.11$, 95% CI [0.08, 0.15], $p = 0.15$), and gender ($\beta = 0.07$, 95% CI [-0.03, 0.17], $p = 0.37$) on parent-reported dominance were also nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta < 0.01$, 95% CI [-0.03, 0.03], $p > 0.99$, $\Delta R^2 < 0.01$; Fig. 2b). The $T \times C \times$ gender interaction was also nonsignificant ($\beta = -0.27$, 95% CI

Table 2 Hierarchical Regression Results Predicting Dominance

Step	Variable	B	β	SE $_{\beta}$	CI $_{\beta}$	<i>p</i>	R ²	F (<i>p</i>)
OV: Observer-rater Dominance								
Model 1: Salivary Hormones								
1	Saliva time	<0.01	0.02	<.01	[0.02, 0.02]	.64	.27	54.73 (<.001)
	Leader–Follower Time	-0.59	-0.52	.06	[-0.63, -0.40]	<.001		
2	Gender	<0.01	<0.01	.10	[-0.19, 0.19]	.998	.30	21.13 (<.001)
	T	-0.12	-0.13	.05	[-0.23, -0.03]	.02		
	C	0.10	0.07	.09	[-0.10, 0.24]	.24		
3	T x C	-0.05	-0.12	.08	[-0.28, 0.05]	.55	.30	17.62 (<.001)
4	T x C x Gender	0.31	0.52	.17	[0.19, 0.85]	.07	.32	12.53 (<.001)
Model 2: Hair Hormones								
1	Leader–Follower Time	-0.58	-0.51	.06	[-0.62, -0.40]	<.001	.26	110.80 (<.001)
2	Gender	0.04	0.02	.12	[-0.22, 0.26]	.74	.29	20.00 (<.001)
	T	0.06	0.07	.05	[-0.04, 0.18]	.27		
	C	-0.02	-0.03	.04	[-0.11, 0.06]	.65		
3	T x C	-0.06	-0.15	.04	[-0.23, -0.07]	.14	.29	16.53 (<.001)
4	T x C x Gender	-0.12	-0.27	.10	[-0.46, -0.08]	.20	.31	10.82 (<.001)
OV: AMS Dominance-Parent								
Model 3: Salivary Hormones								
1	Saliva time	<0.01	-0.02	<.01	[-0.02, -0.02]	.72	.03	1.73 (.14)
	Digits Forward	-0.01	-0.05	.01	[-0.08, -0.03]	.45		
	Digits Backward	<0.01	0.01	.01	[-0.01, 0.03]	.93		
	Trails B Time	-0.002	-0.16	<.01	[-0.16, -0.16]	.01		
2	Gender	0.04	0.06	.04	[-0.02, 0.14]	.38	.03	1.00 (.43)
	T	-0.01	-0.05	.02	[-0.09, -0.005]	.54		
	C	0.01	0.02	.04	[-0.05, 0.09]	.79		
3	T x C	<0.01	0.03	.03	[-0.03, 0.10]	.89	.03	0.88 (.54)
4	T x C x Gender	0.01	0.04	.07	[-0.09, 0.18]	.90	.04	0.72 (.72)
Model 4: Hair Hormones								
1	Digits Forward	-0.01	-0.04	.01	[-0.06, -0.02]	.56	.02	2.10 (.10)
	Digits Backward	<0.01	<0.01	.01	[-0.01, 0.02]	.95		
	Trails B Time	-0.001	-0.16	<.01	[-0.16, -0.16]	.02		
2	Gender	0.05	0.07	.05	[-0.03, 0.17]	.37	.03	0.85 (.53)
	T	0.01	0.03	.02	[-0.01, 0.08]	.66		
	C	0.03	0.11	.02	[0.08, 0.15]	.15		
3	T x C	<0.01	<0.01	.02	[-0.03, 0.03]	.998	.03	0.73 (.65)
4	T x C x Gender	-0.05	-0.27	.04	[-0.35, -0.20]	.23	.05	0.88 (.55)
OV: AMS Dominance-Youth								
Model 5: Salivary Hormones								
1	Saliva Time	<0.01	0.13	<.01	[0.13, 0.13]	.02	.07	5.62 (<.001)
	Digits Forward	<0.01	0.03	.01	[0.02, 0.05]	.61		
	Digits Backward	<0.01	<0.01	.01	[-0.01, 0.02]	.95		

Table 2 (continued)

Step	Variable	B	β	SE $_{\beta}$	CI $_{\beta}$	<i>p</i>	R ²	F (<i>p</i>)
2	Trails B Time	-0.002	-0.23	< .01	[-0.23, -0.23]	< .001		
	Gender	-0.003	-0.01	.04	[-0.08, 0.06]	.93	.08	3.08 (.004)
	T	< 0.01	0.01	.02	[-0.03, 0.05]	.88		
	C	-0.02	-0.05	.03	[-0.11, 0.01]	.45		
3	T x C	0.04	0.29	.03	[0.23, 0.36]	.20	.09	2.91 (.004)
4	T x C x Gender	0.01	0.03	.06	[-0.09, 0.15]	.93	.11	2.66 (.003)
Model 6: Hair Hormones								
1	Digits Forward	0.01	0.05	.01	[0.03, 0.07]	.48	.05	5.49 (.001)
	Digits Backward	-0.0002	-0.001	.01	[-0.02, 0.02]	.99		
	Trails B Time	-0.002	-0.21	< .01	[-0.21, -0.21]	< .001		
2	Gender	0.05	0.09	.04	[0.0005, 0.17]	.22	.06	2.14 (.05)
	T	0.02	0.08	.02	[0.04, 0.12]	.29		
	C	-0.02	-0.11	.02	[-0.14, -0.07]	.16		
3	T x C	0.02	0.18	.02	[0.14, 0.21]	.15	.07	2.15 (.04)
4	T x C x Gender	-0.05	-0.31	.03	[-0.38, -0.25]	.17	.10	2.18 (.02)
OV: MPQ Social Potency-Youth								
Model 7: Salivary Hormones								
1	Saliva Time	< 0.01	0.11	< .01	[0.11, 0.11]	.09	.04	2.33 (.06)
	Digits Forward	-0.01	-0.04	.02	[-0.08, -0.01]	.57		
	Digits Backward	< 0.01	< 0.01	.02	[-0.03, 0.03]	.97		
	Trails B Time	-0.003	-0.18	< .01	[-0.18, -0.17]	.01		
2	Gender	-0.02	-0.02	.07	[-0.16, 0.11]	.73	.10	2.60 (.01)
	T	0.07	0.17	.04	[0.10, 0.24]	.04		
	C	-0.16	-0.21	.06	[-0.33, -0.09]	.01		
3	T x C	0.10	0.44	.06	[0.33, 0.55]	.10	.11	2.65 (.01)
4	T x C x Gender	-0.04	-0.14	.12	[-0.38, 0.09]	.75	.15	2.80 (.002)
Model 8: Hair Hormones								
1	Digits Forward	-0.01	-0.03	.02	[-0.06, 0.001]	.70	.03	2.10 (.10)
	Digits Backward	< 0.01	< 0.01	.02	[-0.03, 0.03]	.97		
	Trails B Time	-0.001	-0.17	< .01	[-0.17, -0.16]	.02		
2	Gender	0.09	0.10	.08	[-0.06, 0.25]	.24	.04	1.06 (.39)
	T	0.04	0.10	.04	[0.03, 0.17]	.23		
	C	0.01	0.02	.03	[-0.04, 0.07]	.82		
3	T x C	0.05	0.25	.03	[0.19, 0.31]	.10	.06	1.30 (.26)
4	T x C x Gender	-0.003	-0.01	.06	[-0.14, 0.11]	.97	.10	1.72 (.08)

In Step 3, models included all main effects and in Step 4, models included all main effects and lower-order two-way interactions. Full model results can be found on the OSF page for this project (<https://osf.io/9n8gf/>)

SE Standard Error, CI Confidence Interval, OV Outcome Variable, T Testosterone, C Cortisol, AMS Achievement Motivation Scale, MPQ Multidimensional Personality Questionnaire

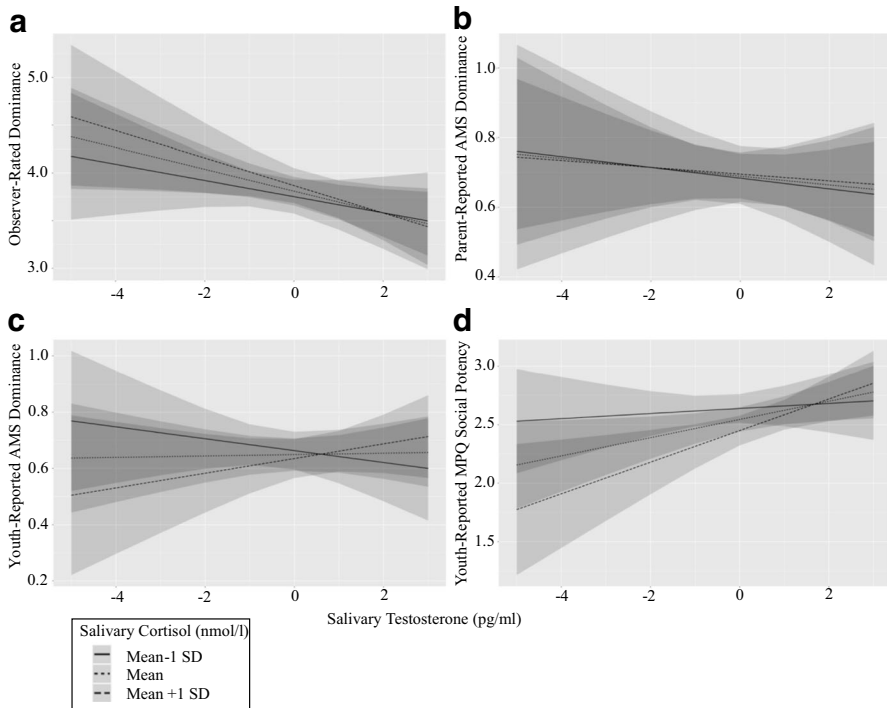


Fig. 1 Salivary testosterone and salivary cortisol do not jointly relate to adolescent dominance across four measures. Intercepts and slopes from multiple regression models were used to plot the relationship between salivary testosterone and adolescent dominance at mean, mean + 1 standard deviation, and mean – 1 standard deviation salivary cortisol levels

[-0.35, -0.20], $p=0.23$, $\Delta R^2=0.02$), though the $T \times C$ interaction was positive in males ($\beta=0.21$, 95% CI [0.14, 0.28], $p=0.36$) and negative in females ($\beta=-0.07$, 95% CI [-0.11, -0.02], $p=0.69$).

Youth-Reported AMS Dominance

For salivary hormones, main effects of T ($\beta=0.01$, 95% CI [-0.03, 0.05], $p=0.88$), C ($\beta=-0.05$, 95% CI [-0.11, 0.01], $p=0.45$), and gender ($\beta=-0.01$, 95% CI [-0.08, 0.06], $p=0.93$) on youth-reported dominance were nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta=0.29$, 95% CI [0.23, 0.36], $p=0.20$, $\Delta R^2=0.01$; Fig. 1c) and the direction of this effect was equivalent in males and females. The $T \times C \times$ gender interaction was also nonsignificant ($\beta=0.03$, 95% CI [-0.09, 0.15], $p=0.93$, $\Delta R^2=0.02$).

For hair hormones, main effects of T ($\beta=0.08$, 95% CI [0.04, 0.12], $p=0.29$), C ($\beta=-0.11$, 95% CI [-0.14, -0.07], $p=0.16$), and gender ($\beta=0.09$, 95% CI [-0.01, 0.17], $p=0.22$) on youth-reported dominance were also nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta=0.18$, 95% CI [0.14, 0.21], $p=0.15$, $\Delta R^2=0.01$; Fig. 2c) and the direction of this effect was equivalent in males and females. The

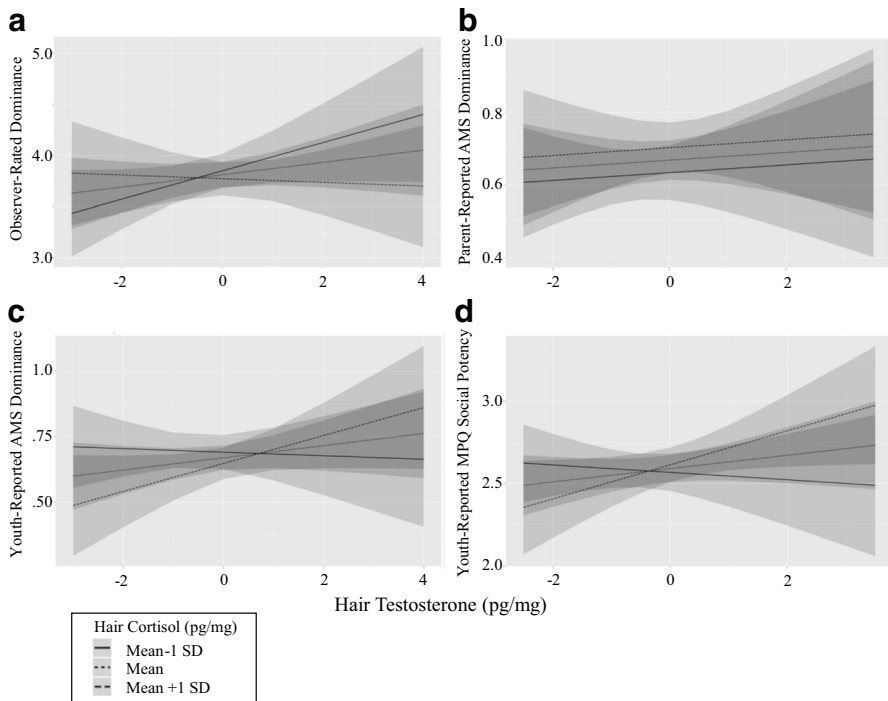


Fig. 2 Hair testosterone and hair cortisol do not jointly relate to adolescent dominance across four measures. Intercepts and slopes from multiple regression models were used to plot the relationship between hair testosterone and adolescent dominance at mean, mean +1 standard deviation, and mean – 1 standard deviation hair cortisol levels

$T \times C \times \text{gender}$ interaction was also nonsignificant ($\beta = -0.31$, 95% CI [-0.38, -0.25], $p = 0.17$, $\Delta R^2 = 0.03$).

Youth-Reported MPQ Social Potency

For salivary hormones, there was a small main effect of T on youth-reported social potency, such that adolescents with higher levels of salivary T were higher in social potency, on average ($\beta = 0.17$, 95% CI [0.10, 0.24], $p = 0.04$). There was also a small main effect of C, such that adolescents with lower levels of salivary C were higher in social potency, on average ($\beta = -0.21$, 95% CI [-0.33, -0.09], $p = 0.01$). The main effect of gender on social potency was nonsignificant ($\beta = -0.02$, 95% CI [-0.16, 0.11], $p = 0.73$). The $T \times C$ interaction was nonsignificant ($\beta = 0.44$, 95% CI [0.33, 0.55], $p = 0.10$, $\Delta R^2 = 0.01$; Fig. 1d) and the direction of this effect was equivalent in males and females. The $T \times C \times \text{gender}$ interaction was also nonsignificant ($\beta = -0.14$, 95% CI [-0.38, 0.09], $p = 0.75$, $\Delta R^2 = 0.04$).

For hair hormones, main effects of T ($\beta = 0.10$, 95% CI [0.03, 0.17], $p = 0.23$), C ($\beta = 0.02$, 95% CI [-0.04, 0.07], $p = 0.82$) and gender ($\beta = 0.10$, 95% CI [-0.06, 0.25],

$p=0.24$) on youth-reported social potency were nonsignificant. The $T \times C$ interaction was nonsignificant ($\beta=0.25$, 95% CI [0.19, 0.31], $p=0.10$, $\Delta R^2=0.02$; Fig. 2d) and the direction of this effect was equivalent in males and females. The $T \times C \times \text{gender}$ interaction was also nonsignificant ($\beta=-0.01$, 95% CI [-0.14, 0.11], $p=0.97$, $\Delta R^2=0.05$).

Robustness Analyses

Robustness checks were conducted by examining analyses in the following conditions: 1) removing all covariates; 2) using non-log-transformed C and T values; 3) excluding participants whose hormone secretion levels may be inflated or decreased due to dairy consumption prior to their visit ($n=64$), use of hormonal birth control ($n=11$), or use of diabetes medication ($n=1$); 4) excluding participants whose salivary T and C values differed substantially from T1 to T2 collection times (i.e., >3 standard deviations from the average T1 to T2 difference; $n=5$); 5) use of wake time rather than saliva sample time as a covariate (for salivary hormone models only); 6) inclusion of pubertal status as an additional covariate; 7) use of executive function tasks rather than leadership task response time as a cognitive performance covariate in models where observer-rated dominance was the dependent variable; and 8) excluding participants whose salivary T values were below the lower limit of detectability ($n=9$). Results of these robustness analyses can be found in [Supplementary Table 6](#).

Overall, the results of our main analyses presented above remained consistent across these robustness tests. In just a few cases, robustness analyses indicated variability in effect sizes that sometimes resulted in a change in statistical significance of effects. For example, the negative main effect of salivary T on observer-rated dominance was no longer statistically significant when covariates were removed, when using non-log transformed T and C values, and when executive function tasks were used as a cognitive performance covariate rather than response time on the leadership task. The positive main effect of salivary T on MPQ Social Potency was no longer statistically significant when covariates were removed, when excluding participants whose T or C values changed greatly between collection times, when pubertal status was included as a covariate, and when removing participants who had consumed dairy on the morning of their visit or who were taking hormonal birth control or diabetes medication. The negative main effect of salivary C on MPQ Social Potency was no longer statistically significant when using non-log-transformed T and C values.

In three cases, statistical significance emerged in robustness analyses in the context of initial null results. First, a significant salivary $T \times C \times \text{gender}$ interaction emerged for observer-rated dominance when non-log-transformed T and C values were used. Second, there was a positive main effect of hair C on parent-reported dominance when non-log-transformed T and C values were used. Third, a significant salivary $T \times C$ interaction emerged for youth-reported dominance when non-log-transformed T and C values were used. However, we interpret these statistically significant effects as likely spurious because they were only present under highly

selective conditions. Further, given the number of tests conducted, there was a high likelihood of observing several spurious positive results. Indeed, all statistically significant effects which emerged in the context of initial null results were reduced to nonsignificance when the Holm (1979) method was used to correct p -values for multiple comparisons within each hypothesis family (e.g., all tests examining salivary hormones and observer-rated dominance; [Supplementary Table 6](#)).

Discussion

Extending previous research on the DHH, we evaluated whether T and C jointly relate to adolescent dominance using task-based, parent-report, and youth self-report measures of adolescent dominance and two measures of hormones (saliva and hair). Evidence for main effects of either T or C on dominance were largely null. The only outcome variable where salivary T and C main effects emerged in typically expected directions was for Social Potency, a subscale of the well-validated MPQ. These single main effects emerging in the expected directions were small but statistically significant, consistent with recent meta-analytic findings (Dekkers et al., 2019). The positive T main effect was variable across robustness analyses, whereas the negative C main effect remained statistically significant across robustness analyses, except when non-log-transformed T and C values were used.

Across methods and hormone measures, evidence in support of the DHH was weak. Point estimates of $T \times C$ interactions were suggestive of a small DHH-consistent effect for observer-rated dominance for both salivary and hair hormones, consistent with the interactions reported in Mehta and Josephs (2010) Study 1 and meta-analytic findings (Dekkers et al., 2019). However, point estimates of the interaction effects were near zero for parent-reported dominance and positive for youth-reported dominance and social potency, highlighting weak support for the DHH across dominance measures. Finally, whereas three-way $T \times C \times \text{gender}$ interactions across methods and hormone measures were nonsignificant, gender-specific analyses indicated the direction of the $T \times C$ interaction was not always equivalent in males and females. This lack of consistent support for the DHH – both for the task-based dominance measure and the three questionnaire measures – may be interpreted in a broader context, which we discuss below.

One interpretation of these findings is that development may play a critical role in hormone-behavior relationships. The current study was conducted with an adolescent sample, and the largely null results suggest the possibility that effects consistent with the DHH may not emerge until after hormone variability associated with puberty has stabilized. A preponderance of studies on the DHH previously have been carried out in young adult samples. Though Dekkers and colleagues (2019) did not find evidence that age moderated the effect in their meta-analysis, this may have been due to restricted age range in the studies included (i.e., only 3 of the 33 studies were conducted with a predominantly adolescent sample). Importantly, within our adolescent sample, the results of the present study did not change when pubertal status was included as a covariate. In addition, two early studies of the DHH involved adolescent boys and found evidence for an effect, but it is important to note these

studies were carried out in justice-involved samples: Dabbs and colleagues examined a sample of late-adolescent offenders (1991), and Popma et al. examined boys referred to a delinquency diversion program (2007). A more recent mixed-gender study of externalizing behaviors in a community sample of adolescents found evidence in support of the DHH for aggression, however (Grotzinger et al., 2018). Thus, the fact that the current study finds no evidence of an effect may be explained by the less clinically severe nature of the community sample, divergent operationalizations of status-seeking (i.e., aggression or violence in the aforementioned studies versus trait and situational dominance in the present study), or a combination of these factors.

Another interpretation of these findings implicates the variable nature of potential moderators across studies, including the clinical status of the population of interest. Indeed, research on hormone-behavior relationships has increasingly indicated that variability in hormone-behavior relationships across studies may be driven by influential moderators (Geniole et al., 2019; Knight et al., 2020). Viewing the DHH literature in its entirety, a cluster of potential moderators emerge (gender, age, clinical status, disinhibitory personality traits; see Knight et al., 2020). These proposed moderators may serve to illuminate meaningful evidence for the DHH by effectively stratifying broader samples in such a way that amplifies the signal of the effect. For example, given that women manifest lower T levels than men, on average, the fact that evidence for the effect is more common in male samples may be an illustration of the extent to which restricted range of variance in female samples serves to suppress the effect overall, making it more difficult to detect. Consistent with this interpretation, meta-analytic evidence suggests that the $T \times C$ interaction effect is statistically significant in men, but not women, though effect sizes for men and women did not significantly differ (Dekkers et al., 2019). Similar patterns of undetected effects might be expected in community (vs. clinical) samples, children and adolescents (vs. adults), and high self-control personality traits (vs. low) – essentially, any time we have reason to expect that restricted range of variance in hormones may make it more difficult to detect an effect, even when it does exist. Not all potential moderators should automatically be dismissed as suspect or evidence of less-than-rigorous practices (e.g., cherry-picking to produce statistically significant findings); however, researchers need to denote theoretically and methodologically justifiable moderators by carefully pre-registering and transparently reporting all analyses, covariates, and hypothesized moderators.

A final interpretation of the present findings is that prior results showing a $T \times C$ interaction in predicting dominance may have been inflated due to the presence of methodological and reporting issues. This interpretation is in line with the findings of Dekkers and colleagues (2019) that some studies in the DHH literature have relied on insufficient sample sizes and likely employed flexible analytic practices. Further, publication bias and selective reporting of measures may have led to the suppression of null findings for $T \times C$ interactions in predicting dominance in the literature (Franco et al., 2014; Simmons et al., 2011). As was illustrated through the results of our robustness analyses, a significant $T \times C$ interaction may be found when a high number of tests are applied across multiple measures of the target construct, but this effect may not generalize across the majority of tests. In the presence of

problematic scientific practices and reporting standards, it is virtually impossible to formulate a certainty or confidence level around the positive empirical support the DHH has garnered in the literature.

Limitations and Future Directions

As discussed extensively by Prasad, Lassetter, and colleagues (2019), several factors may impact hormone measurement validity, including matrix interference (e.g., variance in freeze–thaw cycles and centrifugation and freezer storage periods) and the performance of antibodies used to isolate the substrate being analyzed, which can vary across immunoassay kits. These factors, along with use of different methods to measure hair and salivary hormones (Welker et al., 2016) and saliva collection at only two time points on a single day (Zhang et al., 2018), could affect precision in hormone measurement. Measurement error reduces statistical power even in the context of adequate sample size, and it is common to find largely nonsignificant effects in underpowered studies (Blake & Gangestad, 2020). This raises the possibility that the mostly nonsignificant findings in the present study may be a result of Type II errors secondary to limitations in hormone measurement. Measurement error also contributes to increased Type I error rates, however, particularly in the presence of publication bias. Moreover, while a strength of our study is its in-depth pre-registration, we caution researchers against viewing pre-registration as a panacea. The goal of pre-registrations is to guard against “over-confidence in post hoc explanations” (Nosek et al., 2018, p. 2600) and to increase transparency in the reporting of analyses, ultimately promoting scientific reproducibility.

In the context of these limitations, we offer three suggestions. First, we encourage readers to be critical consumers of scientific products and to view science as a systematic process of uncertainty reduction. As such, we view our project as one of many relevant contributions to the dual-hormone literature. Second, we recommend that pre-registrations using hormonal data include plans for several hormone measurement reliability and validity checks, including inter- and intra-assay coefficients of variability, checks for sex differences or diurnal fluctuations in hormone levels if these are expected, and (if possible) use of multiple methods to measure at least a subset of hormone samples. Many useful checks are outlined in existing papers (e.g., Schultheiss & Stanton, 2009). We recommend that hormone validity checks be conducted in blinded datasets (i.e., datasets that do not include outcome variables) to reduce the potential for bias.

Third, future research using more rigorous methods and practices including increased statistical power and precision, larger sample sizes, pre-registration of hypotheses, measures, and analyses, and improved measurement is still needed. All forms of pre-registration control for analytic flexibility and reporting biases, but Registered Reports may also be used to significantly diminish publication bias for null effects (Nosek et al., 2018). Registered Reports have certain advantages for research on hormone-behavior relationships because they provide the opportunity for reviewers to address methodological concerns prior to data collection. Whereas the present study used hormonal and dominance measures selected from a much

larger study, there are some benefits to designing a study with narrower aims in mind (e.g., testing for hormone effects in a restricted set of dominance-relevant outcomes rather than selecting outcomes from a large set of measures), and these are well-suited for Registered Reports.

An ideal future study of the DHH would be a very large Registered Report that employs multiple methods of measuring 1) hormones and 2) outcome variables such as dominance that are pre-specified prior to data collection. For example, comprehensive assessment of hormones may include measuring C and T in saliva, blood serum, and/or urine via multiple methods (e.g., immunoassays and LC–MS/MS; El-Farhan et al., 2017). This kind of study would be positioned to test the boundaries around dual-hormone effects, and to rigorously investigate theoretically informed potential moderators. Further, measurement error is an often-overlooked hindrance to statistical power (Williams & Zimmerman, 1989), and not all measures of hormones (e.g., immunoassays vs. mass spectrometry; Welker et al., 2016) nor outcomes (e.g., Flake et al., 2017) are created equal. Rigorous construct validation results in tests that are more reliable are purer representations of the latent construct of interest, and therefore produce more signal than those that are created “on-the-fly” or on the basis of face validity alone (Cronbach & Meehl, 1955; Grahek et al., [in press](#)).

Conclusions

In a community sample of adolescents, the present study used a multi-method approach to evaluate the DHH in a series of pre-registered hypotheses and analyses. Results were suggestive of inconsistent support for the DHH across task-based and parent- and self-report questionnaire measures of dominance and across salivary and hair hormones. These results are generally consistent with a recent meta-analysis indicating only marginal support for the DHH (Dekkers et al., 2019). These null effects may indicate that empirical support for the DHH is a result of spurious or cherry-picked findings. However, it is also possible that hormone concentrations in typical community samples are too low to pick up on meaningful variance in hormone effects, but that the dual-hormone effect may appear in some population subgroups or under the influence of specific moderators. We urge researchers to continue to interrogate the DHH in the context of theoretically strong moderators and with an eye toward more precise delineation of those boundary conditions under which the DHH might emerge in human samples. Finally, these results underscore the importance of using open science practices when investigating hormone-behavior relationships. Practices such as pre-registration allow for maximum transparency in manipulation of hormone variables, measures chosen, analytical approaches, and robustness checks, and ultimately offer the only solution to discovery of true DHH effects.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40750-021-00167-3>.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

- Adam, E. K., & Kumari, M. (2009). Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*, *34*(10), 1423–1436.
- Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*(3), 319–345.
- 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*(2), 241–261.
- Bakker, M., Hartgerink, C. H., Wicherts, J. M., & van der Maas, H. L. (2016). Researchers' intuitions about power in psychological research. *Psychological Science*, *27*(8), 1069–1077.
- Blake, K. R., & Gangestad, S. (2020). On attenuated interactions, measurement error, and statistical power: Guidelines for social and personality psychologists. *Personality and Social Psychology Bulletin*, *46*(12), 1702–1711.
- Book, A. S., Starzyk, K. B., & Quinsey, V. L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, *6*(6), 579–599.
- Brown, L. L., Tomarken, A. J., Orth, D. N., Loosen, P. T., Kalin, N. H., & Davidson, R. J. (1996). Individual differences in repressive-defensiveness predict basal salivary cortisol levels. *Journal of Personality and Social Psychology*, *70*(2), 362.
- Burnstein, K. L., Maiorino, C. A., Dai, J. L., & Cameron, D. J. (1995). Androgen and glucocorticoid regulation of androgen receptor cDNA expression. *Molecular and Cellular Endocrinology*, *115*(2), 177–186.
- Carré, J. M., & Archer, J. (2018). Testosterone and human behavior: the role of individual and contextual variables. *Current Opinion in Psychology*, *19*, 149–153.
- Carré, J. M., & McCormick, C. M. (2008). Aggressive behavior and change in salivary testosterone concentrations predict willingness to engage in a competitive task. *Hormones and Behavior*, *54*(3), 403–409.
- Cassidy, T., & Lynn, R. (1989). A multifactorial approach to achievement motivation: The development of a comprehensive measure. *Journal of Occupational Psychology*, *62*(4), 301–312.
- Casto, K. V., Hamilton, D. K., & Edwards, D. A. (2019). Testosterone and cortisol interact to predict within-team social status hierarchy among Olympic-level women athletes. *Adaptive Human Behavior and Physiology*, *5*(3), 237–250.
- Chen, S. Y., Wang, J., Yu, G. Q., Liu, W., & Pearce, D. (1997). Androgen and glucocorticoid receptor heterodimer formation a possible mechanism for mutual inhibition of transcriptional activity. *Journal of Biological Chemistry*, *272*(22), 14087–14092.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.
- Dabbs, J. M., Jurkovic, G. J., & Frady, R. L. (1991). Salivary testosterone and cortisol among late adolescent male offenders. *Journal of Abnormal Child Psychology*, *19*(4), 469–478.
- 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.
- Denson, T. F., Spanovic, M., & Miller, N. (2009). Cognitive appraisals and emotions predict cortisol and immune responses: A meta-analysis of acute laboratory social stressors and emotion inductions. *Psychological Bulletin*, *135*(6), 823.
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 35.
- 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.

- El-Farhan, N., Rees, D. A., & Evans, C. (2017). Measuring cortisol in serum, urine and saliva—are our assays good enough? *Annals of Clinical Biochemistry*, *54*(3), 308–322.
- Field, H. P. (2013). Tandem mass spectrometry in hormone measurement. In M. J. Wheeler (Ed.), *Hormone Assays in Biological Fluids*. (pp. 45–74). Humana Press.
- Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378.
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505.
- Gao, W., Stalder, T., Foley, P., Rauh, M., Deng, H., & Kirschbaum, C. (2013). Quantitative analysis of steroid hormones in human hair using a column-switching LC–APCI–MS/MS assay. *Journal of Chromatography B*, *928*, 1–8.
- Geniole, S. N., Busseri, M. A., & McCormick, C. M. (2013). Testosterone dynamics and psychopathic personality traits independently predict antagonistic behavior towards the perceived loser of a competitive interaction. *Hormones and Behavior*, *64*(5), 790–798.
- Geniole, S. N., Carré, J. M., & McCormick, C. M. (2011). State, not trait, neuroendocrine function predicts costly reactive aggression in men after social exclusion and inclusion. *Biological Psychology*, *87*(1), 137–145.
- Geniole, S. N., Procyshyn, T. L., Marley, N., Ortiz, T. L., Bird, B. M., Marcellus, A. L., Welker, K. M., Bonin, P. L., Goldfarb, B., Watson, N. V., & Carré, J. M. (2019). Using a psychopharmacogenetic approach to identify the pathways through which—and the people for whom—testosterone promotes aggression. *Psychological Science*, *30*(4), 481–494.
- Goldsmith, H. H., & Lemery, K. S. (2000). Linking temperamental fearfulness and anxiety symptoms: A behavior–genetic perspective. *Biological Psychiatry*, *48*(12), 1199–1209.
- Grahek, I., Schaller, M., & Tackett, J. L. (in press). Anatomy of a psychological theory: Integrating construct validation and computational modeling methods to advance theorizing. *Perspectives on Psychological Science*.
- 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.
- Grotzinger, A. D., Mann, F. D., Patterson, M. W., Tackett, J. L., Tucker-Drob, E. M., & Harden, K. P. (2018). Hair and salivary testosterone, hair cortisol, and externalizing behaviors in adolescents. *Psychological Science*, *29*(5), 688–699.
- Harden, K. P., Wrzus, C., Luong, G., Grotzinger, A., Bajbouj, M., Rauters, A., Wagner, G. G., & Riediger, M. (2016). Diurnal coupling between testosterone and cortisol from adolescence to older adulthood. *Psychoneuroendocrinology*, *73*, 79–90.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.
- Johnson, E. O., Kamilaris, T. C., Chrousos, G. P., & Gold, P. W. (1992). Mechanisms of stress: A dynamic overview of hormonal and behavioral homeostasis. *Neuroscience & Biobehavioral Reviews*, *16*(2), 115–130.
- Kirschbaum, C., & Hellhammer, D. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology*, *19*(4), 313–333.
- Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The ‘Trier Social Stress Test’—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1–2), 76–81.
- 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*, 104657.
- Kordsmeyer, T. L., & Penke, L. (2019). Effects of male testosterone and its interaction with cortisol on self- and observer-rated personality states in a competitive mating context. *Journal of Research in Personality*, *78*, 76–92.
- Lahey, B. B., Applegate, B., Waldman, I. D., Loft, J. D., Hankin, B. L., & Rick, J. (2004). The structure of child and adolescent psychopathology: Generating new hypotheses. *Journal of Abnormal Psychology*, *113*(3), 358–385.
- Lausen, A., Broering, C., Penke, L., & Schacht, A. (2020). Hormonal and modality specific effects on males’ emotion recognition ability. *Psychoneuroendocrinology*, *119*, 104719.

- Liening, S. H., Stanton, S. J., Saini, E. K., & Schultheiss, O. C. (2010). Salivary testosterone, cortisol, and progesterone: Two-week stability, interhormone correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. *Physiology & Behavior*, *99*(1), 8–16.
- Mehta, P. H., Jones, A. C., & Josephs, R. A. (2008). The social endocrinology of dominance: Basal testosterone predicts cortisol changes and behavior following victory and defeat. *Journal of Personality and Social Psychology*, *94*(6), 1078.
- Mehta, P. H., & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*(5), 684–692.
- Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, *58*(5), 898–906.
- Mehta, P. H., & Prasad, S. (2015). The dual-hormone hypothesis: A brief review and future research agenda. *Current Opinion in Behavioral Sciences*, *3*, 163–168.
- Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88–99.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606.
- Oyegbile, T. O., & Marler, C. A. (2005). Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. *Hormones and Behavior*, *48*(3), 259–267.
- Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment*, *14*(2), 150.
- Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133.
- Pfafftheicher, S. (2017). Illuminating the dual-hormone hypothesis: About chronic dominance and the interaction of cortisol and testosterone. *Aggressive Behavior*, *43*(1), 85–92.
- Pfafftheicher, S., Landhäußer, A., & Keller, J. (2014). Individual differences in antisocial punishment in public goods situations: The interplay of cortisol with testosterone and dominance. *Journal of Behavioral Decision Making*, *27*(4), 340–348.
- Platje, E., Popma, A., Vermeiren, R. R., Doreleijers, T. A., Meeus, W. H., van Lier, P. A., Koot, H. M., Branje, S. J. T., & Jansen, L. M. (2015). Testosterone and cortisol in relation to aggression in a non-clinical sample of boys and girls. *Aggressive Behavior*, *41*(5), 478–487.
- Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., Jansen, L. M. C., van Engeland, H., & Doreleijers, T. A. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological Psychiatry*, *61*(3), 405–411.
- Ponzi, D., Zilioli, S., Mehta, P. H., Maslov, A., & Watson, N. V. (2016). Social network centrality and hormones: The interaction of testosterone and cortisol. *Psychoneuroendocrinology*, *68*, 6–13.
- Prasad, S., Knight, E. L., & Mehta, P. H. (2019). Basal testosterone's relationship with dictator game decision-making depends on cortisol reactivity to acute stress: A dual-hormone perspective on dominant behavior during resource allocation. *Psychoneuroendocrinology*, *101*, 150–159.
- Prasad, S., Lassetter, B., Welker, K. M., & Mehta, P. H. (2019). Unstable correspondence between salivary testosterone measured with enzyme immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*, *109*, 104373.
- Prasad, S., Narayanan, J., Lim, V. K., Koh, G. C., Koh, D. S., & Mehta, P. H. (2017). Preliminary evidence that acute stress moderates basal testosterone's association with retaliatory behavior. *Hormones and Behavior*, *92*, 128–140.
- Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. Reitan Neuropsychology Laboratory.
- Roelofs, K., van Peer, J., Berretty, E., de Jong, P., Spinhoven, P., & Elzinga, B. M. (2009). Hypothalamus–pituitary–adrenal axis hyperresponsiveness is associated with increased social avoidance behavior in social phobia. *Biological Psychiatry*, *65*(4), 336–343.
- Ronay, R., van der Meij, L., Oostrom, J. K., & Pollet, T. V. (2018). No evidence for a relationship between hair testosterone concentrations and 2D: 4D ratio or risk taking. *Frontiers in Behavioral Neuroscience*, *12*, 30.
- Rubinow, D. R., & Schmidt, P. J. (1996). Androgens, brain, and behavior. *The American Journal of Psychiatry*, *153*(8), 974–984.
- Ruttile, P. L., Shirtcliff, E. A., Armstrong, J. M., Klein, M. H., & Essex, M. J. (2015). Neuroendocrine coupling across adolescence and the longitudinal influence of early life stress. *Developmental Psychobiology*, *57*(6), 688–704.

- Sarkar, A., Mehta, P. H., & Josephs, R. A. (2019). The dual-hormone approach to dominance and status-seeking. In O. C. Schultheiss & P. H. Mehta (Eds.), *Routledge International Handbook of Social Neuroendocrinology*. (1st ed., pp. 113–132). Routledge.
- Schild, C., Aung, T., Kordsmeyer, T. L., Cardenas, R. A., Puts, D. A., & Penke, L. (2020). Linking human male vocal parameters to perceptions, body morphology, strength and hormonal profiles in contexts of sexual selection. *Scientific Reports*, *10*(1), 1–16.
- Schulkin, J., Gold, P. W., & McEwen, B. S. (1998). Induction of corticotropin-releasing hormone gene expression by glucocorticoids: Implication for understanding the states of fear and anxiety and allostatic load. *Psychoneuroendocrinology*, *23*(3), 219–243.
- Schultheiss, O.C. & Stanton, S.J. (2009). Assessment of salivary hormones. In E. Harmon-Jones & J.S. Beer (Eds.), *Methods in social neuroscience* (pp. 17–44). Guilford Publications.
- Sherman, G. D., Lerner, J. S., Josephs, R. A., Renshon, J., & Gross, J. J. (2016). The interaction of testosterone and cortisol is associated with attained status in male executives. *Journal of Personality and Social Psychology*, *110*(6), 921.
- Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: correspondence between hormonal and physical development. *Child Development*, *80*(2), 327–337.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.
- Stalder, T., Steudte, S., Miller, R., Skoluda, N., Dettenborn, L., & Kirschbaum, C. (2012). Intraindividual stability of hair cortisol concentrations. *Psychoneuroendocrinology*, *37*(5), 602–610.
- Tackett, J. L., Herzhoff, K., Harden, K. P., Page-Gould, E., & Josephs, R. A. (2014). Personality × hormone interactions in adolescent externalizing psychopathology. *Personality Disorders: Theory, Research, and Treatment*, *5*(3), 235.
- Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*(3), 411.
- Tellegen, A. (1982). Brief manual for the multidimensional personality questionnaire. *Unpublished manuscript, University of Minnesota, Minneapolis*, 1031–1010.
- Tilbrook, A. J., Turner, A. I., & Clarke, I. J. (2000). Effects of stress on reproduction in non-rodent mammals: The role of glucocorticoids and sex differences. *Reviews of Reproduction*, *5*(2), 105–113.
- Van Den Bos, W., Golka, P., Effelsberg, D., & McClure, S. (2013). Pyrrhic victories: The need for social status drives costly competitive behavior. *Frontiers in Neuroscience*, *7*, 189.
- van Honk, J., Schutter, D. J., Hermans, E. J., Putman, P., Tuiten, A., & Koppeschaar, H. (2004). Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. *Psychoneuroendocrinology*, *29*(7), 937–943.
- Viaua, V. (2002). Functional cross-talk between the hypothalamic-pituitary-gonadal and-adrenal axes. *Journal of Neuroendocrinology*, *14*(6), 506–513.
- Wechsler, D. (2008). *Wechsler adult intelligence scale—Fourth Edition (WAIS-IV)*. NCS Pearson.
- Welker, K. M., Lassetter, B., Brandes, C. M., Prasad, S., Koop, D. R., & Mehta, P. H. (2016). A comparison of salivary testosterone measurement using immunoassays and tandem mass spectrometry. *Psychoneuroendocrinology*, *71*, 180–188.
- Welker, K. M., Lozoya, E., Campbell, J. A., Neumann, C. S., & Carré, J. M. (2014). Testosterone, cortisol, and psychopathic traits in men and women. *Physiology & Behavior*, *129*, 230–236.
- Williams, R. H., & Zimmerman, D. W. (1989). Statistical power analysis and reliability of measurement. *The Journal of General Psychology*, *116*(4), 359–369.
- Wingfield, J. C., Hegner, R. E., Dufty, A. M., Jr., & Ball, G. F. (1990). The “challenge hypothesis”: Theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *The American Naturalist*, *136*(6), 829–846.
- Zhang, Q., Chen, Z., Chen, S., Yu, T., Wang, J., Wang, W., & Deng, H. (2018). Correlations of hair level with salivary level in cortisol and cortisone. *Life Sciences*, *193*, 57–63.

Authors and Affiliations

Allison N. Shields¹  · Cassandra M. Brandes¹  · Kathleen W. Reardon²  ·
Raul A. España¹  · Jennifer L. Tackett¹ 

¹ Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208, USA

² Center for Applied Psychological and Family Studies, The Family Institute at Northwestern University, 618 Library Pl, Evanston, IL 60201, USA