Hormones and Behavior 92 (2017) 82–92

Contents lists available at ScienceDirect

Hormones and Behavior

journal homepage: www.elsevier.com/locate/yhbeh

Performance during competition and competition outcome in relation to testosterone and cortisol among women

Andrea Henry [a,b,*], Jason R. Sattizahn [a], Greg J. Norman [a], Sian L. Beilock [a,c], Dario Maestripieri [b,d]

[a] Department of Psychology, The University of Chicago, USA
[b] Institute for Mind and Biology, The University of Chicago, USA
[c] Committee on Education, The University of Chicago, USA
[d] Department of Comparative Human Development, The University of Chicago, USA

ARTICLE INFO

Article history:
Received 8 January 2016
Revised 25 March 2017
Accepted 28 March 2017
Available online 24 May 2017

Keywords:
Testosterone
Cortisol
Competition
Status
Motivation
Performance
Women
Men

ABSTRACT

A contribution to a special issue on Hormones and Human Competition.
This study investigated the relation between competition, testosterone (T), and cortisol (C) in women. One hundred and twenty female participants competed against a male confederate in a computerized laboratory task. The task was preprogrammed so that half the women won and half of the women lost the competition. T and C concentrations were measured in saliva samples collected at four time points before and after the competition. Accuracy and reaction time during the competition were recorded. T and C increased directly after the competition, though not significantly for C, and then decreased over time regardless of the competition outcome. Regression analyses demonstrated that baseline T was significantly and positively associated with competition accuracy, though only in individuals who were low in C. Individuals who were high in C showed no relation between T and accuracy. This relation was further qualified by competition outcome. Losers of the competition showed a significant positive relation between baseline T levels and competition accuracy, though only if they were low in C. No relation was found between T and accuracy in losers who were high in C. Winners of the competition showed no relation between T and accuracy, regardless of whether C levels were high or low. These results are in line with the dual-hormone hypothesis, whereby the effects of T on status-seeking behaviors are dependent on C levels for individuals whose status is threatened.

1. Introduction

Competitive interactions afford individuals the opportunity to gain or maintain high social status, which allows access to important resources related to survival and reproduction in a social species (Buss, 1988; Stockley and Bro-Jørgensen, 2011; West-Eberhard, 1979). Two hormones that have been strongly implicated in competition and behaviors related to gaining and maintaining social status are testosterone (T) and cortisol (C) (Bateup et al., 2002; Booth et al., 1989; Costa and Salvador, 2012; Edwards et al., 2006; Edwards and Kurlander, 2010; Jiménez et al., 2012; Kivlighan et al., 2005; Mazur and Lamb, 1980; Mazur et al., 1992; Mazur et al., 1997; Mehta and Josephs, 2006; Mehta et al., 2009; Mehta et al., 2015; Oliveira et al., 2013; Schultheiss et al., 2005; Stanton and Schultheiss, 2007; Suay et al., 1999; van Anders and Watson, 2007; van der Meij et al., 2010).

Early studies examining the relation between status and T in humans found a link between baseline levels of T, assertiveness and aggressiveness, and other status-seeking behaviors (Cashdan, 1995; Dabbs and Dabbs, 2001; Grant and France, 2001; Mehta et al., 2008; Mehta et al., 2009). Individuals with higher baseline levels of T were rated by their peers to be more socially dominant (Cashdan, 1995), were found to be more assertive (Grant and France, 2001), performed better in competitions where they competed alone as compared to group competitions (Mehta et al., 2009), and also showed a greater inclination to compete again following the competition if they won (Mehta et al., 2008). Although individual differences in baseline T levels may be relatively stable over time and be influenced by genetic characteristics, prenatal hormonal exposure, or stable environmental conditions (Kempenaers et al., 2008), T levels are also known to fluctuate predictably in relation to changes in reproductive condition or the social environment (Wingfield et al., 1990). According to the Biosocial Model of Status (Mazur, 1985), T responses to social challenges should differ among individuals, depending on their previous experiences and dominance position (Carre and Olmstead, 2015; Casto and Edwards, 2016; Hamilton et al., 2015). One assertion of the Biosocial Model of Status is that individuals who have recently won a competition are more likely

* Corresponding author at: Institute for Mind and Biology, The University of Chicago, 940 E. 57th street, Chicago, IL 60637, USA.
E-mail address: andreahenry@uchicago.edu (A. Henry).

http://dx.doi.org/10.1016/j.yhbeh.2017.03.010
0018-506X/

to experience a rapid rise in T, whereas individuals who have recently lost a competition are more likely experience a rapid decline in T. Numerous studies have supported this prediction, although these effects are often weak and have been explored more extensively in men than in women (Booth et al., 1989; Costa and Salvador, 2012; Geniole et al., 2016; Gladue et al., 1989; Jiménez et al., 2012; Mazur and Booth, 1998; Mazur and Lamb, 1980; Mazur et al., 1992; Oliveira et al., 2009; Serrano et al., 2000; van Anders and Watson, 2007).

A second assertion of the Biosocial Model of Status is that higher T, both baseline levels and fluctuations in response to the environment, is associated with greater motivation to compete to gain or maintain social status (e.g., Booth et al., 1989; Costa and Salvador, 2012; Hamilton et al., 2015; Suay et al., 1999), however there is no strong and unequivocal support for this assumption. Some previous studies of competition have used mood as a proxy for motivation (Booth et al., 1989; Mazur and Lamb, 1980), while others have relied on self-report measures of motivation (Costa and Salvador, 2012; Suay et al., 1999). Performance throughout the actual competition may reflect how engaged an individual is and the extent to which he or she wants to win. However, the direct impact of baseline T or fluctuations in T on competition performance itself has been difficult to quantify. In field studies of sports activities, it can be difficult to objectively track an individual's performance throughout a competition, so in some cases subjective measures have been used (Serrano et al., 2000; Trumble et al., 2012). Whenever objective measures of performance have been related to T levels (González-Bono et al., 1999; Kivlighan et al., 2005; Trumble et al., 2012), hormonal fluctuations were likely confounded by the physical nature of the competition (Kraemer and Ratamess, 2005; Webb et al., 1984). Laboratory studies that have quantitatively tracked performance during competition have found mixed results, with T both positively (Costa and Salvador, 2012; Mehta et al., 2009; Schultheiss et al., 2005) and negatively (Kivlighan et al., 2005; Mehta et al., 2009; van Anders and Watson, 2007) associated with performance.

Similar to T, C fluctuates in response to challenging or threatening situations (Del Giudice et al., 2011; Dickerson and Kemeny, 2004; Sapolsky, 2004). C prepares an individual to appropriately respond to a perceived challenge or threat (Del Giudice et al., 2011), and it therefore has been studied in relation to status seeking behaviors and competition (Salvador, 2005; Salvador and Costa, 2009). The studies that have examined the relation between C and competition outcome have produced mixed results, with some studies showing increases in C across both winners and losers, increases in C in either the winners or the losers, or no significant changes in C in response to competition (Bateup et al., 2002; Booth et al., 1989; Costa and Salvador, 2012; Edwards and Kurlander, 2010; Edwards et al., 2006; Elias, 1981; Filaire et al., 2001; Jiménez et al., 2012; Kivlighan et al., 2005; Mazur et al., 1997; Oliveira et al., 2009).

It has been suggested that the relation between T and status seeking behaviors may be best qualified through its interaction with C (Mehta and Josephs, 2010). This dual-hormone hypothesis posits that T is associated with status-seeking or dominant behaviors only when C is low, such that individuals who are high in T and low in C demonstrate greater dominant behaviors than those individuals who are high in T and high in C. When C is high, T is unrelated to status-seeking or dominant behaviors (Mehta and Josephs, 2010; Mehta and Prasad, 2015). In other words, status-seeking or dominant behaviors are jointly regulated by T and C. This relation has been further qualified by environmental context. Support for the dual-hormone hypothesis has been found in men in a competitive setting, but only in those individuals who have recently lost a competition (Mehta and Josephs, 2010). Men who recently lost a competition were more likely to choose to compete again in the same competition if their T levels were high and C levels were low. Men who recently won a competition did not show this relation between their desire to compete again and their T and C levels. In other words, men high in T and low in C chose to compete again only after they lost status through losing the competition (Mehta and Josephs, 2010).

Despite the growing number of studies examining the relation between competition, T, and C, many aspects of the relation between competition, T, and C remain unexplored. First, given the paucity of competition studies involving women (in terms of both women competing against other women and women competing against men), it is not clear whether the hormonal responses to competition observed in men also reliably occur in women. Second, to our knowledge, all studies to date have explored hormonal fluctuations during same-sex competitive interactions. It remains unclear whether hormonal changes observed in same-sex competitions also occur when competing against the opposite sex. Third, the relation between hormones and performance during a competition remains unclear, particularly whether T, C, or their interaction may be associated with performance in the competition and whether this may be qualified by the competition outcome (i.e. winning or losing).

This study was designed to address some of these gaps in our knowledge of the relation between competition, T, and C. Our first goal was to investigate the effects of winning or losing on hormone levels among women who compete against men. We hypothesized that women who won the competition against men would demonstrate an increase in their T levels, whereas women who lost the competition would experience a decline in T. A second goal of this study was to explore in detail the possible relation between hormone levels (baseline and fluctuations), competition outcome, and performance during the competition. We hypothesized that winners would demonstrate overall higher accuracy in the competition than losers due to increased motivation to secure a win after positive feedback on initial rounds of the competition (Vallerand, 1983; van Dijk and Kluger, 2011). Additionally, we hypothesized that women who had higher levels of baseline T or greater increases in T across the competition would demonstrate higher accuracy as they attempted to win the competition. In other words, we hypothesized that increased T levels would be associated with better accuracy in the competition, as T may affect motivation and effort to do well in the competition in order to maintain or gain status against the competitor. In line with the dual-hormone hypothesis, we hypothesized that C would affect competition performance through its interaction with T, such that individuals high in T and low in C would perform the best.

2. Material and methods

2.1. Participants

One hundred and twenty women (age: $M = 21.75$, $SD = 2.96$) were recruited online and through flyers. Women were eligible to participate if they answered >5, the midpoint, on two 9-point Likert scale questions ("I am good at math", "It is important to me that I am good at math"). This eligibility requirement was used as it pertained to a second task that was completed following the competition. Participants were reimbursed 1.5 course credits or $15.00 for study participation. Those participants assigned to the winning condition received an additional $5.00 for participation.

2.2. Study design

Participants engaged in a competition against a male confederate who they were led to believe was another participant. Participants were randomly assigned to a winning or losing condition for the competition task. The asterisks in the competitive paradigm (described below) were presented in a random order in all blocks. Accuracy and reaction time in the competitive paradigm were compared across winners and losers. Saliva samples were collected at various time points throughout the tasks to assess hormone levels via ELISA (see below). All experimental procedures were approved by the Social Science Institutional Review Board of the University of Chicago.

2.3. Procedures

Prior to participation in the study, participants completed an online battery of surveys. The surveys included: demographics (age, sex, race, ethnicity, normal exercise regimen, smoker/non-smoker, etc.), Spielberger's State-Trait Anxiety Inventory – Trait (STAI-T; Spielberger et al., 1983), Abbreviated Math Anxiety Survey (Hopko et al., 2003), Sensitivity to Punishment and Sensitivity to Reward Questionnaire (Torrubia et al., 2001), the Competitiveness Questionnaire (Griffin-Pierson, 1990), the Rosenberg Self-Esteem Scale (Rosenberg, 1965), the MacArthur Scale of Subjective Status (Adler et al., 2000), and four questions from the importance subscale of the Collective Self-Esteem Scale (Luhtanen and Crocker, 1992) modified to assess the perceived importance of gender identity to self-definition. All computerized tasks were programmed using EPrime software (Psychology Software Tools, Inc.). In lab surveys were administered following the computerized tasks via surveymonkey.com. These surveys included questions related to the prior tasks and additional demographic information.

All experimental conditions were run in the afternoon between 12 PM and 5 PM. Prior to the participant's arrival, a Caucasian male confederate (age: 25) was brought into the lab and seated in front of a computer. When the participant arrived, the confederate was already in place at his desk. The participant was walked past the confederate, was told that the confederate was another participant who had arrived slightly early, and was seated at a desk to the right hand side of the confederate. The confederate and the participant were separated from each other by a folding room separator. A consent form was then administered and signed. After a 5-min resting period, the first saliva sample was collected from the participant to serve as a baseline measure for hormone levels. This sample was collected approximately 10 min after arrival to the lab.

Following the resting period, the participant was notified via computerized instructions that she and the other ‘participant’ would be competing against each other in a task. The participant was not aware that the winner or loser of the competition was predetermined. It was then explained that the competition would consist of numerous rounds of a visual motor response time task (see below) that would be scored both on accuracy and response time. The participant learned that the winner would receive an additional $5.00 for his or her participation, which was included to increase investment in the task. At that point in time the participant completed 8 practice trials. After the practice trials, a pre-task survey question was administered regarding how well the participant thought she would do on the task, ranging from very poor to very well on a 5-point Likert scale. The participant then waited for the experimenter who walked over to the computers and emphasized that the task would take place in real time between the participants and that they should try their best. After this the task began.

A modified version of the Serial Response Task (Schultheiss et al., 2005; Wirth et al., 2006) was used as the competitive paradigm. A problem encountered by previous studies of competition in women was their possible lack of investment in the competitive task (Costa and Salvador, 2012; Mazur et al., 1997). Specifically, laboratory competition paradigms have used tasks similar in nature to video games, which women may not be interested or invested in. This problem could be particularly significant when investigating women who compete against men. To address this problem and increase motivation to engage in the competition, we used a gender-neutral competition in which neither males nor females would hold an unequal interest or a competitive edge. Unlike previous laboratory studies that used a rigged pencil-and-paper tracing task that does not easily allow for minute differences in performance to be tracked, the computerized competitive task allowed for mapping an individual's behavioral performance in the competition, specifically accuracy and reaction time, to hormonal measures.

The overall competition lasted 10 min and consisted of ten rounds of a visual motor response task. In each round, asterisks were presented randomly at four locations on the computer screen. Each of the four asterisk locations was mapped to a certain letter on the keyboard. The participants were told to press the key assigned to the asterisk location. When the key was pressed, the asterisk disappeared and moved to a new location on the screen. Participants then pressed the key assigned to the new asterisk location. Slightly modified from the original version of the task, in our task the asterisks moved randomly, as opposed to sequentially, to new locations on the screen for the entirety of the round. Our asterisks were also presented toward the outer four corners of the screen rather than in the center of the screen. Participants were informed that after each round of the competition the computer would calculate their score based on response time and accuracy and then compare the scores of the competitors to determine and announce a winner for the round. Each round began with a screen announcing the round number, after which a countdown occurred. The participants then performed the response task for 50 s, followed by a black screen that said, “Calculating and comparing scores…” for 2 s. When a participant won a round, the next screen displayed a green background for 2 s that said “You have won this round” and was accompanied by the sound of a jingling bell. When a participant lost a round, the screen displayed a red background for 2 s that said “You have lost this round” and was accompanied by a buzzing sound. The final screen of the round was blank but maintained the background color of the feedback screen (either green or red) for 3 s. Participants who ‘won’ the competition won all rounds except the second and fifth round. Participants who ‘lost’ the competition lost all rounds except the second and fifth round. We did not want the winners to win all rounds and the losers to lose all rounds to avoid giving the participants the impression that the outcome of the competition was rigged, and consequently reduce their effort on the task. Accuracy and response time were recorded for each round throughout the competition, despite the preprogrammed ‘win’ or ‘loss’ of the competition.

Following the competition, the winner was verbally congratulated for his or her win by the experimenter and was reminded of the extra $5.00 compensation he or she would receive at the end of the experiment. Immediately afterward, the second saliva sample was collected, which was approximately 10 min after the start of the competition. Participants then filled out Spielberger's State-Trait Anxiety Inventory – State (STAI-S; Spielberger et al., 1983) and the PANAS-X (Watson and Clark, 1994). Affect was measured as a manipulation check and as a proxy for engagement in the competition. Participants also answered various questions regarding the competition, such as, “Which factors contributed to your final result: my own capacity, my effort, luck, task difficulty, opponent's ability; Use the following areas to gauge task enjoyment: interesting, enjoyable, fun, boring, and a waste of time; Characterize the task with regard to: effort, frustration, performance, stress, and perceived difficulty”. Following the online surveys, participants provided a third saliva sample. This collection occurred approximately 10 to 15 min following the end of the competition. Following the competition, participants completed another cognitive task that lasted approximately 10 min, and then the fourth saliva sample was collected. Lastly, the participants were debriefed.

2.4. Saliva sample collection and hormonal assays

Samples were collected between 12:00 pm and 5:00 pm by a trained experimenter. Throughout the experiment, four saliva samples were collected via passive drool into a plastic tube, and nothing was given to participants to help stimulate salivation. After saliva collection samples were frozen at —20C until later assayed in house in Dr. Dario Maestripieri's Behavioral Biology Laboratory. On the day of processing, they were thawed and centrifuged at 3*g* for 15 min. Each of the four samples was then assayed using an ELISA kit provided by Salimetrics for both T and C. The interassay coefficient of variation for T was

6.37%, and the interassay coefficient of variation for C was 6.12%. The intraassay coefficient of variation for T was 3.98%, and the intraassay coefficient of variation for C was 6.08%.

In a reminder email sent one day prior to participation, participants were instructed to refrain from eating and drinking and from consuming any caffeine products for 2 h prior to their arrival in the lab. Information was collected after the competition about the following variables that may potentially influence hormone concentrations: time at which the participants woke up that morning, whether they consumed caffeine that morning, the time they had consumed caffeine that morning, the time they last ate that morning, whether they consumed tobacco products that morning, whether they had consumed alcohol within the past 24 h, whether they had exercised within the past 24 h, how far along they were in their menstrual cycle, how long their cycle typically lasted, their average amount of exercise per week, whether they smoke, whether they were on medication that influence Glucocorticoids (as specified by examples: steroids, antiglucocorticoids, antidepressants, etc.), their arrival time at the lab, their first saliva sampling relative to the time they woke up, their relationship status, their relationship length or how long they had been in a relationship, and whether or not they were on contraceptives (as specified by examples: oral contraceptives; IUD; Birth control patch; Nuva Ring; Hormonal injections).

Baseline T levels were not normally distributed, demonstrating a positive skew, and therefore all T values were log transformed. Log transformation resulted in a normal distribution. Log transformed T values were therefore used for all T analyses. Of the 120 participants in the competition, two women reported not cycling for nine months or more due to contraceptive use. These participants were excluded from data analyses. Additionally, two other participants were more than three standard deviations from the baseline log transformed T mean and were also excluded as outliers from the analyses. Baseline levels of C were not normally distributed and were therefore log transformed, which resulted in a normal distribution. Therefore, all C analyses were conducted with log transformed data. One participant had a baseline C level that was three standard deviations above the mean and was excluded as an outlier from the analyses. Overall, after excluding outliers, 115 subjects were used in the final analyses.

2.5. Statistical procedures

When comparing the changes in T and C across the competition between winners and losers, repeated measures ANOVAs were used with competition outcome as the between subjects variable. When comparing the changes in accuracy and reaction time across the competition between winners and losers, repeated measures ANOVAs were used with competition outcome as the between subjects variable. To determine how pre-competition T, pre-competition C, and competition outcome affected performance during the competition, a hierarchical multiple linear regression was run to test for the three conditional effects of T, C, and competition outcome and all possible interactions on competition accuracy. Prior to running the regression, all variables were standardized or centered. Simple slopes analyses were performed to further qualify the significant interactions found in the regression analyses.

3. Results

There were no significant differences between winners and losers in demographic variables (age, race, or ethnicity), psychological/experiential variables (interpersonal competitiveness, goal competitiveness, subjective social status ratings, trait anxiety, self-esteem, depression, fear of negative evaluation from peers, whether they currently played sports, whether they ever played sports, or how much an individual identified with being female), or any variables that may influence hormone concentrations (see Methods).

3.1. Competition and affect

To determine whether participants were invested in the competitive task and to perform a manipulation check, affect and subjective evaluation of the competition were compared between winners and losers. Losers reported significantly more anxiety following the competition than winners. As compared to winners, losers also reported significantly more negative affect, more hostility, more guilt, more sadness, less positive affect, less joviality, less self-assurance, and less attentiveness. As compared to losers, winners evaluated the competition as more interesting, more enjoyable, more fun, more difficult, and less boring. Means, standard deviations, *p*-values, and effect sizes are reported in Table 1. Overall, participants demonstrated reliable mood changes following competition, with increases in positive mood states in the winners and increases in negative mood states in the losers, which suggests that participants were impacted by the competition and invested in the task.

3.2. Sources of individual differences in baseline T and C levels

There were no significant differences in baseline T levels between winners (M = 81.47, SD = 30.15) and losers (M = 80.98, SD = 26.75): $t_{113} = -0.12$, $p = 0.91$. Across all 115 participants for whom T was analyzed, baseline T levels were significantly lower in individuals who were on contraceptives ($n = 30$) (M = 65.92, SD = 24.92) than in those who were not on contraceptives ($n = 85$) (M = 86.63, SD = 27.68): $t_{113} = -4.27$, $p < 0.001$, $d = 0.93$. Baseline T levels were also significantly lower in participants who were in a relationship ($n = 57$) (M = 72.80, SD = 24.61) than in those who were not in a relationship ($n = 58$) (M = 89.51, SD = 29.60): $t_{113} = -3.48$, $p < 0.001$, $d = 0.62$, and baseline levels of T were especially low in those participants who reported being in a relationship for six months or longer ($n = 47$) (M = 69.15, SD = 23.42) as compared to those not in a long-term relationship ($n = 68$) (M = 89.57, SD = 28.67): $t_{113} = -4.40$, $p < 0.001$, $d = 0.78$. To determine which of these three variables was driving the effect on baseline T levels, a linear regression was run in which contraceptive use, relationship status, and relationship length were each entered as dichotomous predictor variables. Baseline log T values served as the dependent variable. The overall model was significant: $F_{3,114} = 8.06$, $p < 0.001$ with an adjusted R square value of 0.16. The only predictor that was a significant predictor of baseline T levels after controlling for the other two was contraceptive use: $\beta = 0.225$, $t = 2.10$, $p = 0.038$.

No significant differences were detected in baseline levels of C between winners (M = 0.22, SD = 0.11) and losers (M = 0.25, SD = 0.17): $t_{113} = -0.88$, $p = 0.38$. Across all study participants for whom C was analyzed, baseline C levels were lower in participants who were

Table 1
Self-reported anxiety, affect and subjective evaluation of competition after completing the competition.

	Winners		Losers		T-value	*p*-Value	Cohen's d
	Mean	SD	Mean	SD			
State anxiety	36.55	8.96	44.39	9.73	−4.49	<0.001	0.84
Negative affect	13.93	5.32	16.25	4.79	−2.45	0.016	0.46
Hostility	6.98	1.74	8.96	3.90	−3.51	0.001	0.66
Guilt	7.31	2.87	9.89	4.83	−3.48	0.001	0.65
Sadness	6.71	3.00	8.82	3.44	−3.53	0.001	0.65
Positive affect	27.0	6.82	22.39	7.37	3.49	0.001	0.65
Joviality	20.03	6.61	15.74	6.05	3.637	<0.001	0.68
Self-assurance	14.53	4.62	11.81	4.75	3.12	0.002	0.58
Attentiveness	13.16	2.91	11.25	3.35	3.27	0.001	0.61
Interesting	2.38	0.77	2.11	0.67	2.04	0.044	0.37
Enjoyable	2.45	0.78	2.05	0.81	2.67	0.009	0.50
Fun	2.47	0.78	2.05	0.83	2.749	0.007	0.52
Difficult	3.36	0.79	2.75	0.71	4.331	<0.001	0.81
Boring	1.83	0.68	2.19	0.92	−2.43	0.017	0.45

in a relationship (M = 0.20, SD = 0.10) than in those who were not in a relationship (M = 0.26, SD = 0.17): $t_{113} = -2.51$, $p = 0.014$, $d = 0.46$, and baseline levels of C were especially low in those participants who reported being in a relationship for six months or longer (M = 0.18, SD = 0.08) as compared to those not in a long-term relationship (M = 0.27, SD = 0.16): $t_{113} = -3.80$, $p < 0.001$, $d = 0.75$. To determine which of these variables was driving the effect on baseline log C levels, a linear regression was run in which the two dichotomous predictor variables of relationship status and relationship length were entered into the model. Baseline log C values served as the dependent variable. The overall model was significant: $F_{2,114} = 7.84$, $p = 0.001$ with an adjusted R square value of 0.11. The predictor that was a significant predictor of baseline C levels after controlling for the other predictor was relationship length: $\beta = 0.49$, $t = 2.99$, $p = 0.003$.

Due to the effects of contraceptives on baseline levels of T, contraceptive use was statistically controlled for in all analyses of T data. Due to the effects of relationship length on baseline levels of C, relationship length was statistically controlled for in all analyses of C data. Specifically, when running a repeated measures ANOVA to examine hormonal fluctuations over time, either contraceptive use or relationship length was entered as a covariate, depending on which hormone was analyzed. In the regression model where both T and C were analyzed, both contraceptive use and relationship length were entered as covariates.

3.3. T changes after the competition in winner and losers

Across all participants, a repeated measures, Greenhouse-Geisser corrected ANOVA using Log T at the four time points as the within-subjects factor and 'win'/'lose' condition as the between-subjects factor revealed a significant main effect of time: $F_{2.49,\ 112} = 5.96$, $p = 0.001$, $\eta^2 = 0.051$. Post-hoc, pairwise comparisons indicated that T levels were significantly higher at: time point one than at time point four ($p < 0.001$), time point two than at time point one ($p = 0.048$), time point two than at time point three ($p < 0.001$), time point two than at time point four ($p < 0.001$), and at time point three than at time point four ($p = 0.001$). There was no significant difference in T values from time point one to time point three. There was no main effect of winning/losing: $F_{1,112} = 0.01$, $p = 0.92$, and no significant interaction between time and winners/losers: $F_{2.49,\ 112} = 0.48$, $p = 0.70$, indicating that winners and losers did not differ significantly in their changes in T over time (See Fig. 1). Overall, T levels increased from time point 1 to time point 2 (i.e. raw T increased by 2.8% from time point 1 to time point 2) and then declined from time point two to time points three and four. There were no significant differences in T changes between winners and losers. The means and standard deviations of raw T values are reported in Table 2.

3.4. C changes after the competition in winner and losers

Across all study participants, Greenhouse-Geisser corrected, repeated measures ANOVA with log C at the four time points as the within-subjects and 'win'/'lose' as the between-subjects variable revealed a main effect of time that approached significance: $F_{2.34,\ 112}$: 2.36, $p = 0.087$, $\eta^2 = 0.021$. Post-hoc, pairwise comparisons demonstrated that C levels were significantly higher at: time point one than at time point four ($p = 0.002$), time point two than at time point four ($p < 0.001$), and at time point three than at time point four ($p < 0.001$). The difference between C levels at time point one and time point two approached significant ($p = 0.075$). There was no significant main effect of winning and losing: $F_{1,\ 112} = 0.25$, $p = 0.62$, and no significant interaction between winning and losing and change in C over time: $F_{2.34,\ 112} = 1.66$, $p = 0.19$ (See Fig. 1). Overall, C levels declined over time, though not significantly, and there were no significant differences in C changes between winners and losers. The means and standard deviations of raw C values are reported in Table 3.

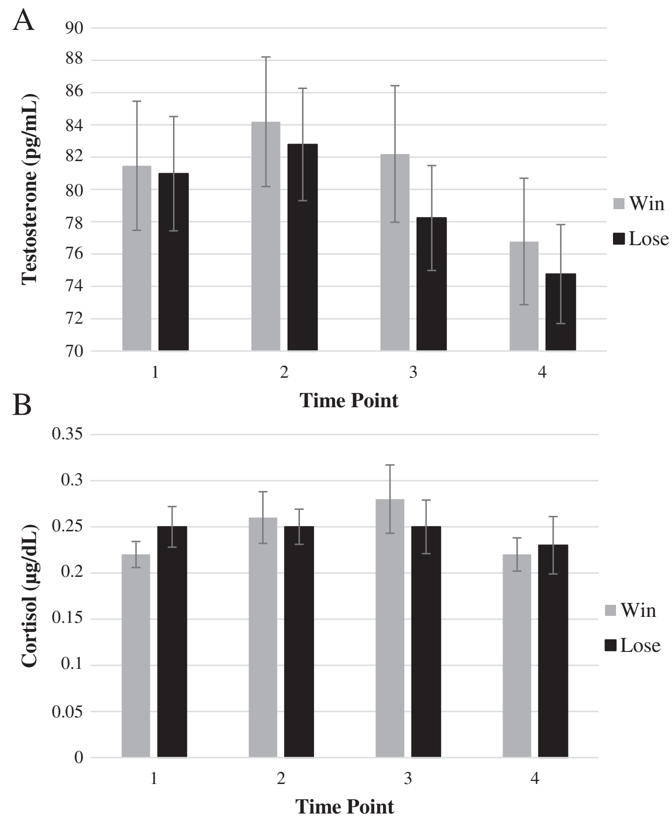

Fig. 1. A. Mean (±SEM) of raw testosterone concentrations at 4 time points in winners ($n = 58$) and losers ($n = 57$). Time point 1 occurred approximately 10 min after arrival into the lab. Time point 2 occurred directly after the competition ended, which was approximately 10 min after the competition began. Time point 3 occurred approximately 10 min after the competition ended. Time point 4 occurred approximately 10 min after the third time point. B. Mean (±SEM) raw cortisol concentrations at 4 time points in winners ($n = 58$) and losers ($n = 57$). Time point 1 occurred approximately 10 min after arrival into the lab. Time point 2 occurred directly after the competition ended, which was approximately 10 min after the competition began. Time point 3 occurred approximately 10 min after the competition ended. Time point 4 occurred approximately 10 min after the third time point.

3.5. Accuracy and reaction time during the competitive task in winners and losers

Overall, both winners and losers performed above 90% accuracy across all rounds, indicating both that the task was relatively easy and, importantly, that all study participants were interested in the task enough to perform above chance. There were no significant differences in accuracy ($t_{113} = 1.00$, $p = 0.32$) (winners: M = 97.44, SD = 2.94, and losers: M = 96.82, SD = 3.62) or reaction time ($t_{113} = -0.05$, $p = 0.96$) (winners: M = 361.51, SD = 38.56, and losers: M = 361.94, SD = 62.64) on the first round of the competition between winners and losers, suggesting that the initial involvement in the competition prior

Table 2
Means and standard deviations of raw testosterone concentrations (pg/mL) at 4 time points in both winners ($n = 58$) and losers ($n = 57$).

	Winners		Losers	
	Mean	SD	Mean	SD
Time point 1	81.47	30.15	80.98	26.75
Time point 2	84.19	30.56	82.79	26.24
Time point 3	82.2	32.18	78.24	24.57
Time point 4	76.78	29.82	74.76	23.11

Table 3
Means and standard deviations of raw cortisol concentrations (μg/dL) at 4 time points in both winners ($n = 58$) and losers ($n = 57$).

	Winners		Losers	
	Mean	SD	Mean	SD
Time point 1	0.22	0.11	0.25	0.17
Time point 2	0.26	0.22	0.25	0.14
Time point 3	0.28	0.28	0.25	0.22
Time point 4	0.22	0.14	0.23	0.23

to feedback was similar across winners and losers. A repeated measures ANOVA with a Greenhouse-Geisser correction revealed a main effect of round number on accuracy: $F_{2.47,\ 113} = 11.71$, $p < 0.001$, $\eta^2 = 0.094$, as well as on reaction time: $F_{4.73,\ 113} = 93.94$, $p < 0.001$, $\eta^2 = 0.454$: accuracy decreased while reaction time increased across the rounds for both winners and losers. For accuracy, there was also a significant main effect of winning and losing: $F_{1,113} = 8.24$, $p = 0.005$, $\eta^2 = 0.068$, such that winners had a significantly higher overall accuracy across all ten rounds (M = 96.23, SD = 2.81) than losers (M = 94.19, SD = 4.62). There was no significant interaction between round number and winning and losing on accuracy: $F_{2.47,\ 113} = 1.34$, $p = 0.26$, though as one can see from the graph, losers decreased in accuracy across the rounds more so than winners. No main effect of winning and losing on reaction time was found: $F_{1,113} = 0.13$, $p = 0.72$ and no interaction between round number and winning and losing was found with regard to reaction time: $F_{4.73,\ 113} = 1.42$, $p = 0.22$ (See Fig. 2). The means and standard deviations for average competition accuracy and reaction time across each of the rounds of the competition are reported in Table 4.

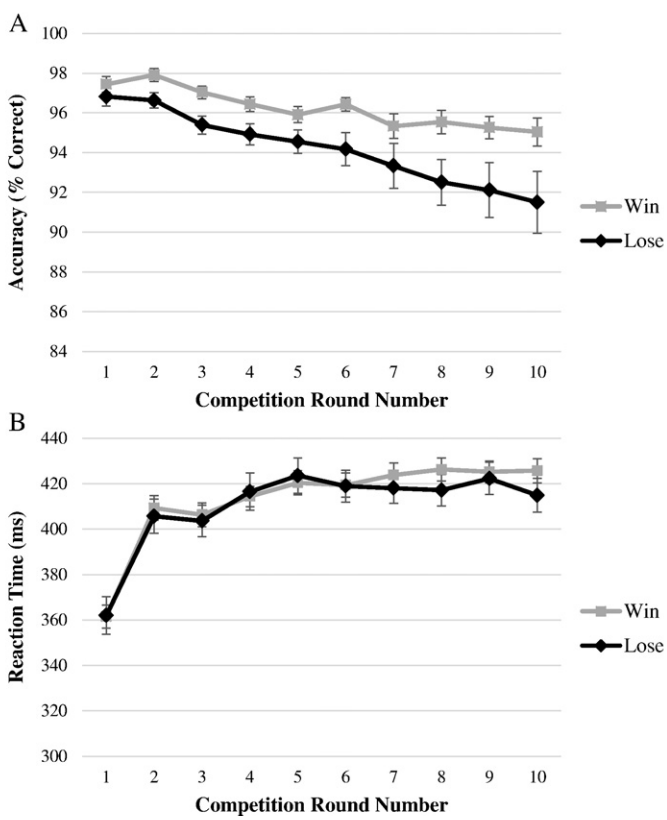

Fig. 2. A. Accuracy (mean ± SEM) across the 10 rounds of the competition task in winners ($n = 58$) and losers ($n = 57$). B. Reaction time (mean ± SEM) across the 10 rounds of the competition task in winners ($n = 58$) and losers ($n = 57$).

Table 4
Means and standard deviations of average competition accuracy (percent correct) and reaction time (milliseconds) in each of the ten rounds of competition in both winners ($n = 58$) and losers ($n = 57$).

	Competition accuracy				Competition reaction time			
	Winners		Losers		Winners		Losers	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Round 1	97.44	2.94	96.82	3.62	361.51	38.56	361.94	62.64
Round 2	97.91	2.49	96.63	2.87	409.32	41.13	405.63	57.15
Round 3	97.04	2.46	95.39	3.43	406.31	39.74	403.64	52.43
Round 4	96.44	2.78	94.92	3.99	414.37	34.03	416.49	62.16
Round 5	95.91	3.04	94.55	4.43	420.4	40.67	423.55	58.04
Round 6	96.43	2.48	94.17	6.26	419.37	40.48	418.96	53.16
Round 7	95.33	4.74	93.33	8.52	423.81	40.71	418.02	50.00
Round 8	95.54	4.51	92.51	8.65	426.26	38.47	417.12	51.81
Round 9	95.26	4.32	92.11	10.44	425.2	36.77	422.27	53.76
Round 10	95.04	5.30	91.5	11.78	425.72	40.38	414.88	56.22

3.6. Competition performance (accuracy), baseline T, and baseline C

Across all study participants, log T values at time points one, two, three, and four were significantly positively correlated with competition accuracy averaged across all ten rounds (T1: $r = 0.21$, $p = 0.025$, $r^2 = 0.044$, T2: $r = 0.20$, $p = 0.043$, $r^2 = 0.04$, T3: $r = 0.20$, $p = 0.031$, $r^2 = 0.04$, T4: $r = 0.24$, $p = 0.011$, $r^2 = 0.063$). Thus, higher T levels were associated with higher competition accuracy. Changes in T across the competition were not associated with competition accuracy (all p-values > 0.65). Furthermore, log T values were not correlated with any average reaction time measures (all p-values > 0.48). Similar to T, log C values at time points three and four were positively associated with average competition accuracy across all ten rounds (T3: $r = 0.19$, $p = 0.042$, $r^2 = 0.036$, T4: $r = 0.22$, $p = 0.022$, $r^2 = 0.048$, and a similar but nonsignificant relation existed for log C values at time points one and two: T1: $r = 0.15$, $p = 0.118$; T2: $r = 0.17$, $p = 0.074$). Thus, higher C was associated with higher competition accuracy. Changes in C across the competition were not associated with competition accuracy (all p-values > 0.25). Additionally, log C values were not correlated with average reaction time (all p-values > 0.55).

To test whether the effect of baseline T on accuracy during the competition may be moderated by baseline C, a hierarchical multiple linear regression was run in which average competition accuracy was entered as the dependent variable, and baseline log T, baseline log C, competition outcome (coded as −1 for losers and +1 for winners in all regression analyses), contraceptive use and relationship length were entered as predictor variables. In the first model, these variables were entered as predictors without any interaction terms. This overall model was significant [adjusted $R^2 = 0.097$, $F_{5,114}$: 3.44, $p = 0.006$]. Competition outcome significantly predicted competition accuracy ($\beta = 0.28$, $t = 3.13$, $p = 0.002$), such that winners had higher average accuracy than losers. All other predictors were not significantly related to competition accuracy (Table 5a). Each of the possible two way interactions were added into the second model so that it included all single predictor variables as well as all possible two way interactions. This overall model was significant [adjusted $R^2 = 0.17$, $F_{8,114}$: 3.89, $p = 0.001$]. Competition outcome again significantly predicted competition accuracy ($\beta = 0.2$, $t = 3.12$, $p = 0.002$), such that winners had higher average accuracy than losers. Additionally, the two-way interaction between baseline T and baseline C was significant: $\beta = -0.28$, $t = -3.17$, $p = 0.002$. All other predictors were not significantly related to competition accuracy (Table 5b). The three way interaction was added into the third model, which resulted in a model that included all single predictor variables, each of the two-way interactions, and the three-way interaction. This overall model was significant [adjusted $R^2 = 0.26$, $F_{9,114}$: 5.34, $p < 0.001$]. The two-way interaction between baseline T and baseline C was significant: $\beta = -0.27$, $t = -3.21$, $p = 0.002$. Additionally, the

Table 5
Regression models examining the relation between competition accuracy and baseline T, baseline C, competition outcome, and their interactions. Covariates include contraceptive use and relationship length.

a: Model 1 – All single predictors.

Coefficients	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	0.95	0.004		234.012	0.00
Competition outcome	0.011	0.004	0.28	3.13	0.002
Testosterone	0.007	0.004	0.17	1.53	0.13
Cortisol	0.004	0.004	0.097	0.93	0.36
Contraceptive use	0.002	0.005	0.037	0.32	0.75
Relationship length	−0.003	0.005	−0.087	−0.74	0.46

b: Model 2 – All single predictors and all two way interactions.

Coefficients	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	0.96	0.004		221.31	0.00
Competition outcome	0.011	0.003	0.27	3.12	0.002
Testosterone	0.007	0.004	0.17	1.58	0.12
Cortisol	0.004	0.004	0.095	0.92	0.36
Contraceptive use	0.005	0.005	0.11	0.97	0.33
Relationship length	−0.004	0.005	−0.11	−0.93	0.35
T × Competition outcome	−0.004	0.004	−0.091	−0.91	0.37
C × Competition outcome	−0.001	0.004	−0.031	−0.31	0.76
T × C	−0.010	0.003	−0.28	−3.17	0.002

c: Model 3 – All single predictors, all two way interactions, and the three way interaction.

Coefficients	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	0.958	0.004		234.02	0.00
Competition outcome	0.005	0.004	0.14	1.52	0.13
Testosterone	0.006	0.004	0.14	1.39	0.17
Cortisol	0.004	0.004	0.091	0.94	0.35
Contraceptive use	0.004	0.005	0.087	0.82	0.42
Relationship length	−0.004	0.004	−0.095	−0.87	0.39
T × Competition outcome	−0.002	0.004	−0.045	−0.47	0.64
C × Competition outcome	−0.002	0.004	−0.058	−0.61	0.55
T × C	−0.010	0.003	−0.27	−3.21	0.002
T × C × Competition outcome	0.011	0.003	0.330	3.68	0.000

three-way interaction between baseline T, baseline C, and competition outcome was significant: β = 0.33, t = 3.68, p < 0.001. All other effects were not significant (Table 5c).

After detecting a significant two-way interaction and a significant three-way interaction, a simple slopes analysis was performed to test the simple effects of both interactions. For the two-way interaction, the effects of T on average accuracy were examined at high and low levels of cortisol. A significant and positive effect of T on competition accuracy was found for participants who were low in C (one standard deviation below the mean) (B = 0.38, t = 3.15, p = 0.002), such that individuals who were low in C increased in average competition accuracy as T increased. There was no significant effect of T on competition accuracy for individuals who were high in C (one standard deviation above the mean) (B = −0.10, t = −0.79, p = 0.43). (See Fig. 3)

For the three-way interaction, the effect of T on average accuracy in the competition was examined at high and low levels of cortisol in both winners and losers. In winners, there was no significant effect of T on competition accuracy for individuals who were either low in C (one standard deviation below the mean) (B = 0.12, t = 0.59, p = 0.56) or for individuals who were high in C (one standard deviation above the mean) (B = 0.20, t = 0.93, p = 0.36). In losers, there was a significant effect of T on competition accuracy for individuals who were low in C (one standard deviation below the mean) (B = 0.55, t = 3.33, p = 0.002), such that those individuals who were low in C increased in competition accuracy as T increased. However, there was not a significant effect of T on competition accuracy for losers who were high in C (one standard deviation above the mean) (B = −0.28, t = −1.48, p = 0.15) (See Fig. 4). The effects of T and C on competition accuracy were driven by individuals in the losing condition, as there were no effects of T on competition accuracy for either high or low C individuals in the winning condition. Competition accuracy values were negatively skewed. Importantly, after transforming accuracy with an arcsine transformation, the aforementioned associations remain between competition accuracy and baseline T, baseline C, and competition outcome.

It should be noted two participants demonstrated competition accuracy scores that were three standard deviations below the overall group mean for competition accuracy. To reduce the influence of these outliers

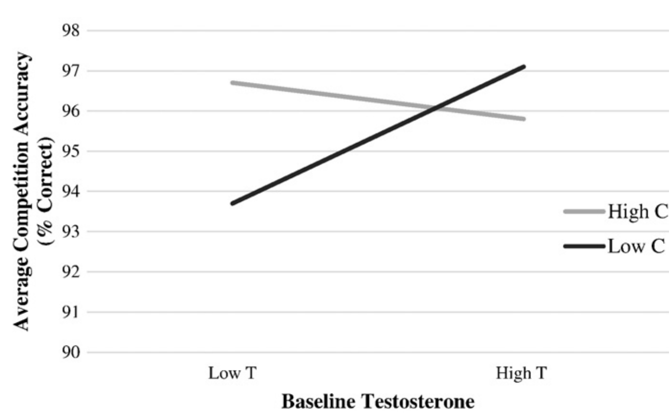

Fig. 3. Average competition accuracy across both winners (n = 58) and losers (n = 57) as a function of baseline testosterone and cortisol levels. High = one standard deviation above the mean; low = one standard deviation below the mean.

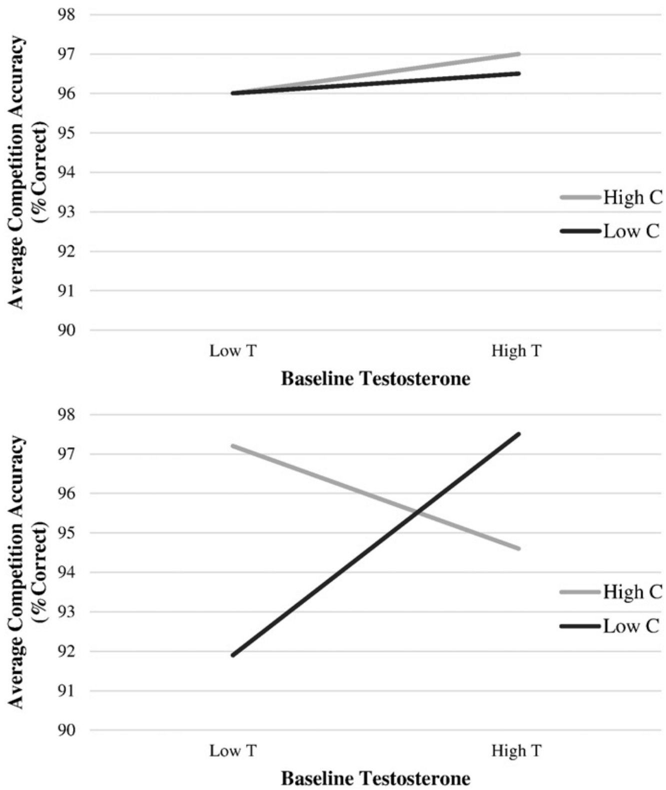

Fig. 4. A. Average competition accuracy in winners ($n = 58$) as a function of baseline testosterone and cortisol levels. High = one standard deviation above the mean; low = one standard deviation below the mean. B. Average competition accuracy in losers ($n = 57$) as a function of baseline testosterone and cortisol levels. High = one standard deviation above the mean; low = one standard deviation below the mean.

on the aforementioned results, the two data points were winsorized, such that they were replaced with the average competition accuracy that was three standard deviations below the mean. A similar hierarchical multiple regression was then run in which accuracy scores that included the winsorized values were entered as the dependent variable, and baseline log T, baseline log C, competition outcome, contraceptive use and relationship length were entered as predictor variables. The first model included each of these variables as predictors without any interaction terms. This overall model was significant [adjusted $R^2 = 0.085$, $F_{5,114}$: 3.13, $p = 0.011$]. Competition outcome significantly predicted competition accuracy ($\beta = 0.30$, $t = 3.24$, $p = 0.002$), such that winners had higher average accuracy than losers. All other predictors were not significantly related to competition accuracy. In the second model, each of the single predictors was included in the model along with each of the possible two way interactions. This overall model was significant [adjusted $R^2 = 0.095$, $F_{8,114}$: 2.50, $p = 0.016$]. Competition outcome again significantly predicted competition accuracy ($\beta = 0.29$, $t = 3.17$, $p = 0.002$), such that winners had higher average accuracy than losers. The two-way interaction between baseline T and baseline C approached significance: $\beta = -0.16$, $t = -1.77$, $p = 0.080$. All other predictors were not significantly related to competition accuracy. The third model included all single predictor variables, each of the two way-interactions, and the three-way interaction. This overall model was significant [adjusted $R^2 = 0.13$, $F_{9,114}$: 2.90, $p = 0.005$]. The two-way interaction between baseline T and baseline C approached significance: $\beta = -0.15$, $t = -1.72$, $p = 0.089$. Additionally, the three-way interaction between baseline T, baseline C, and competition outcome was significant: $\beta = 0.22$, $t = 2.24$, $p = 0.027$. All other effects were not significant. These results suggest that these two participants were not driving the observed relation between baseline T, baseline C, competition outcome, and competition performance.

4. Discussion

Our study is unique in that it is one of the few experimental investigations in which women directly competed against a man in a laboratory task and also in which multiple saliva samples were taken to effectively track hormonal fluctuations in response to competition. Winning or losing a competitive task against a man was associated with significant increases in women's post-competition concentrations of T relative to baseline, whereas C did not significantly differ post-competition from baseline. Both T and C generally declined over time (presumably reflecting the normal diurnal fluctuations in these hormones or some general psychophysiological response to the testing situation) without significant differences between winners and losers. Some previous studies in which women competed against other women reported a clear increase in T in winners and a decrease in T in losers (Costa and Salvador, 2012; Jiménez et al., 2012; Oliveira et al., 2009), while others reported an increase in T in both winners and losers (Bateup et al., 2002; Edwards and Kurlander, 2010; Edwards and O'Neal, 2009; Edwards et al., 2006; Hamilton et al., 2009), or no effect of competition on T (Mazur et al., 1997). These mixed results in hormonal fluctuations in response to competition may be due to differences in methodologies across studies. Athletic competitions are typically associated with robust increases in both T and C in both winners and losers (Casto and Edwards, 2016), which may contribute to the mixed findings. It is possible that for studies completed in the laboratory the mixed results in hormonal responses to competition may be due to differences in the number of samples collected and to the timing of sample collection surrounding the competition. The number of samples taken in the present study and the short time period between each sample allowed for minute changes in hormone levels to be tracked over time in response to competition. The present study suggests that there is in fact a rise in T levels in response to competition that may best be captured directly following the competitive interaction. It is possible that prior studies that have not found a hormonal response to competition or have found inconsistent results (in either women or men) have sampled saliva too long after the competition has occurred to reliably detect hormonal changes due to context. Future studies may benefit from more frequent samples obtained sooner after the competitive interaction.

Despite our failure to show a differential change in T in response to competition between winners and losers, we did find evidence to support one assumption of the Biosocial Model of Status: there is an association between baseline T and status-related motivation in situations in which status is threatened. Baseline T was related to increased competition accuracy, which could serve as a proxy for motivation to secure higher status in competition. Specifically, higher levels of baseline T were associated with higher competition accuracy. Importantly, this relation between high baseline T and increased accuracy was further qualified by an individual's baseline C levels. Women whose status was threatened by losing the competition showed a significant relation between baseline T, baseline C and competition accuracy, such that high T was associated with higher accuracy only when C levels were low. The relation between high T and accuracy was blocked when C levels were high. Importantly, this relation only existed in individuals who whose status was threatened by losing the competition and who received negative feedback about their status throughout the competition. Despite the fact that T was associated with contraceptive use, which is in line with previous research (Schultheiss et al., 2005; Edwards and O'Neal, 2009), contraceptive use was not driving the associations found between T and competition accuracy as it was statistically controlled for in all analyses. Due to the ease and simplicity of the competitive task, the differences in competition accuracy were not large between winners and losers, with all participants scoring on average above 90% accuracy. Despite the small amount of variance in performance across participants, the behavioral measures still captured significant differences between winners and losers that were associated with individual differences in hormone levels. It is possible that more robust findings in line with the current results may be observed in competitive interactions that are more difficult or taxing, as the variance in performance may be greater.

The association between T, C and competition performance are directly in line with the dual-hormone hypothesis and prior findings that align with the dual-hormone hypothesis. The relation between baseline T and status-seeking or dominant behaviors is qualified by baseline C levels. Individuals with high T and low C are rated by their peers to be more dominant (Mehta and Josephs, 2010) and engage in greater dominance related behaviors (Popma et al., 2007; Pfattheicher et al., 2014). Furthermore, in a competitive setting the relation between dominance behaviors, baseline T and baseline C levels is moderated by the competition outcome. Only when an individual's status is threatened, or when an individual loses the competitive interaction, are baseline T and C levels related to an individual's desire to choose to compete again in order to potentially regain status (Mehta and Josephs, 2010). The current results are in line with and extend prior findings. In the current study, we extend support for the dual-hormone hypothesis to women in a competition in a laboratory setting. Furthermore, rather than rely on post-competition subjective measures of desire to compete again, behavioral measures of accuracy captured their motivation to secure high status in the competition in real time and this behavioral measure was related to baseline hormone levels. Previous research has demonstrated a relation between baseline T and competition accuracy (Mehta et al., 2009), changes in T and competition accuracy on a second, subsequent day of competition (Zilioli and Watson, 2014), and changes in T in men and implicit learning during a competition (Schultheiss et al., 2005). Whereas T has previously been associated with competition performance measures, to our knowledge, this is the first time the dual-hormone hypothesis has been supported with behavioral performance measures collected during a competition between two individuals.

The mechanism underlying the relation between T, C, and competition performance in losers may be related to the functional link between the hypothalamic-pituitary-gonadal (HPG) axis and the hypothalamic-pituitary-adrenal (HPA) axis. Glucocorticoids and Corticotropin Releasing Hormone (CRH) released by the HPA axis inhibit functionality of the HPG axis at numerous levels. Increased HPA axis activity has been associated with inhibition of Gonadotropin-Releasing Hormone release (GnRH), reduced responsiveness to GnRH in the pituitary, and decreased sex steroid output through direct effects on the gonads (Johnson et al., 1992). In tissues where both androgen receptors and glucocorticoid receptors are expressed, increases in glucocorticoids are associated with downregulation of androgen receptor expression (Burnstein et al., 1995), making these tissues less sensitive to circulating levels of androgens and ultimately inhibiting the effects of T on these target tissues. As such, high levels of C have been shown to reduce circulating T levels and inhibit or block the cellular effects of T. Behavioral studies on the dual-hormone hypothesis provide evidence that high levels of C block the behavioral effects of T (Mehta and Prasad, 2015), ultimately resulting in a null association between high levels of T and status-seeking behaviors when C levels are high.

Another possible mechanism by which C and T may interact to affect competition performance or status-seeking behavior requires a different level of analysis, moving from a level of analysis on a cellular level toward a neural circuitry level of analysis. In the present study, T and C were found to interact to affect competition performance only in losers who experienced a threat to their status. Previous research has shown activation of brain regions associated with threat, particularly the amygdala, hypothalamus, and the periaqueductal gray, in response to angry faces; moreover, increased activation in these brain regions was related to the exogenous increases in T (Goetz et al., 2014). Another fMRI study conducted with female subjects demonstrated that the amygdala and hypothalamus increased in activation to cues of social threat, and this response was particularly enhanced when baseline T levels were high and baseline C levels were low (Hermans et al., 2008). Given these findings, T and C may be affecting neurotransmission within the amygdala and hypothalamus when an individual's status within the competition is threatened, which in turn impacts behavioral performance. High levels of T and low levels of C may enhance threat processing, whereas threat processing may be disrupted in individuals with high levels of T and high levels of C. Enhanced threat processing may result in an increased motivation to respond to the status threat and protect an individual's status.

One limitation of the current study is the reliance on Enzyme Immunoassay (EIA) to estimate T levels in saliva. There is a measurement error associated with EIAs whereby T levels on the lower range of measurement are inflated (i.e. T levels of women). This inflation of T estimates is likely due to cross-reactivity with other steroids. Furthermore, a lack of concordance exists between EIA estimates of T and more accurate techniques for estimating T levels, such as liquid chromatography tandem mass spectrometry (Welker et al., 2016). As the current study utilized an EIA to estimate T levels in women, it is possible that measurement error masked any potential differences that may exist between winners and losers in response to the competition. In addition to taking more frequent samples to determine hormonal fluctuations in response to competition, future studies may benefit from using more accurate and sensitive measures for estimating T levels, such as liquid chromatography tandem mass spectrometry (Welker et al., 2016).

Another limitation of the current study is the correlational nature of our findings. The associations found here between baseline T, C, and competition performance in losers provide novel support for the dual-hormone hypothesis. However, future studies that examine the relation between competition performance and hormone levels should administer exogenous T to better determine the causal effects of increased T and decreased C on competition performance. Furthermore, the mechanisms responsible for the relation between baseline hormones and competition performance that are outlined above are speculative at best. To better determine the neural networks associated with increased performance in a competitive interaction and how these may relate to hormone levels, future fMRI studies could exogenously administer T and explore changes in brain activity while an individual competes to determine which brain regions are associated with increased or decreased performance in real time and how this performance relates to hormone levels.

Funding statement

This study was funded by intramural funds from the University of Chicago (2-90927). The authors declare no conflicts of interest regarding funding or publication.

References

Adler, N.E., Epel, E.S., Castellazzo, G., Ickovics, J.R., 2000. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol. 19:586–592. http://dx.doi.org/10.1037/0278-6133.19.6.586.

Bateup, H.S., Booth, A., Shirtcliff, E.A., Granger, D.A., 2002. Testosterone, cortisol, and women's competition. Evol. Hum. Behav. 23:181–192. http://dx.doi.org/10.1016/S1090-5138(01)00100-3.

Booth, A., Shelley, G., Mazur, A., Tharp, G., Kittok, R., 1989. Testosterone, and winning and losing in human competition. Horm. Behav. 23:556–571. http://dx.doi.org/10.1016/0018-506X(89)90042-1.

Burnstein, K.L., Maiorino, C.A., Dai, J.L., Cameron, D.J., 1995. Androgen and glucocorticoid regulation of androgen receptor cDNA expression. Mol. Cell. Endocrinol. 115, 177–186.

Buss, D.M., 1988. The evolution of human intrasexual competition: tactics of mate attraction. J. Pers. Soc. Psychol. 54, 616–628.

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

Cashdan, E., 1995. Hormones, sex, and status in women. Horm. Behav. 37, 354–366.

Casto, K.V., Edwards, D.A., 2016. Testosterone, cortisol, and human competition. Horm. Behav. 82, 21–37.

Costa, R., Salvador, A., 2012. Associations between success and failure in a face-to-face competition and psychobiological parameters in young women. Psychoneuroendocrinology 37:1780–1790. http://dx.doi.org/10.1016/j.psyneuen.2012.03.012.

Dabbs, J.M., Dabbs, M.G., 2001. Heroes, Rogues, and Lovers: Testosterone and Behavior. McGraw-Hill, New York.

Del Giudice, M., Ellis, B.J., Shirtcliff, E.A., 2011. The adaptive calibration model of stress responsivity. Neurosci. Biobehav. Rev. 35:1562–1592. http://dx.doi.org/10.1016/j.neubiorev.2010.11.007.

Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol. Bull. 130:355–391. http://dx.doi.org/10.1037/0033-2909.130.3.355.

Edwards, D.A., Kurlander, L.S., 2010. Women's intercollegiate volleyball and tennis: effects of warm-up, competition, and practice on saliva levels of cortisol and testosterone. Horm. Behav. 58:606–613. http://dx.doi.org/10.1016/j.yhbeh.2010.06.015.

Edwards, D.A., O'Neal, J.L., 2009. Oral contraceptives decrease saliva testosterone but do not affect the rise in testosterone associated with athletic competition. Horm. Behav. 56:195–198. http://dx.doi.org/10.1016/j.yhbeh.2009.01.008.

Edwards, D.A., Wetzel, K., Wyner, D.R., 2006. Intercollegiate soccer: saliva cortisol and testosterone are elevated during competition, and testosterone is related to status and social connectedness with teammates. Physiol. Behav. 87:135–143. http://dx.doi.org/10.1016/j.physbeh.2005.09.007.

Elias, M., 1981. Serum cortisol, testosterone, and testosterone-binding globulin responses to competitive fighting in human males. Aggress. Behav. 7:215–224. http://dx.doi.org/10.1002/1098-2337(1981)7:3<215::AID-AB2480070305>3.0.CO;2-M.

Filaire, E., Maso, F., Sagnol, M., Ferrand, C., Lac, G., 2001. Anxiety, hormonal responses, and coping during a judo competition. Aggress. Behav. 27:55–63. http://dx.doi.org/10.1002/1098-2337(20010101/31)27:1<55::AID-AB5>3.0.CO;2-H.

Geniole, S.N., Bird, B.M., Ruddick, E.L., Carre, J.M., 2016. Effects of Competition Outcome on Testosterone Concentrations in Humans: An Updated Meta-Analysis. Horm, Behav.

Gladue, B.A., Boechler, M., McCaul, K.D., 1989. Hormonal response to competition in human males. Aggress. Behav. 15:409–422. http://dx.doi.org/10.1002/1098-2337(1989)15:6<409::AID-AB2480150602>3.0.CO;2-P.

Goetz, S.M.M., Tang, L., Thomason, M.E., Diamond, M.P., Hariri, A.R., Carré, J.M., 2014. Testosterone rapidly increases neural reactivity to threat in healthy men: a novel two-step pharmacological challenge paradigm. Biol. Psychiatry 76, 324–331.

González-Bono, E., Salvador, A., Serrano, M.A., Ricarte, J., 1999. Testosterone, cortisol, and mood in a sports team competition. Horm. Behav. 35:55–62. http://dx.doi.org/10.1006/hbeh.1998.1496.

Grant, V.J., France, J.T., 2001. Dominance and testosterone in women. Biol. Psychol. 58: 41–47. http://dx.doi.org/10.1016/S0301-0511(01)00100-4.

Griffin-Pierson, S., 1990. The competitiveness questionnaire: a measure of two components of competitiveness. Meas. Eval. Couns. Dev. 23, 108–115.

Hamilton, L.D., Carré, J.M., Mehta, P.H., Olmstead, N., Whitaker, J.D., 2015. Social neuroendocrinology of status: a review and future directions. Adapt. Hum. Behav. Physiol. 1: 202–230. http://dx.doi.org/10.1007/s40750-015-0025-5.

Hamilton, L.D., van Anders, S.M., Cox, D.N., Watson, N.V., 2009. The effect of competition on salivary testosterone in elite female athletes. Int. J. Sports Physiol. Perform. 4, 538–542.

Hermans, E.J., Ramsey, N.F., van Honk, J., 2008. Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. Biol. Psychiatry 63:263–270. http://dx.doi.org/10.1016/j.biopsych.2007.05.013.

Hopko, D.R., Mahadevan, R., Bare, R.L., Hunt, M.K., 2003. The abbreviated math anxiety scale (AMAS): construction, validity, and reliability. Assessment 10:178–182. http://dx.doi.org/10.1177/1073191103252351.

Jiménez, M., Aguilar, R., Alvero-Cruz, J.R., 2012. Effects of victory and defeat on testosterone and cortisol response to competition: evidence for same response patterns in men and women. Psychoneuroendocrinology 37:1577–1581. http://dx.doi.org/10.1016/j.psyneuen.2012.02.011.

Johnson, E.O., Kamilaris, T.C., Chrousos, G.P., Gold, P.W., 1992. Mechanisms of stress: a dyanmic overview of hormonal and behavioral homeostasis. Neurosci. Biobehav. Rev. 16, 115–130.

Kempenaers, B., Peters, A., Foerster, K., 2008. Sources of individual variation in plasma testosterone levels. Philos. Trans. R. Soc. B 363:1711–1723. http://dx.doi.org/10.1098/rstb.2007.0001.

Kivlighan, K.T., Granger, D.A., Booth, A., 2005. Gender differences in testosterone and cortisol response to competition. Psychoneuroendocrinology 30:58–71. http://dx.doi.org/10.1016/j.psyneuen.2004.05.009.

Kraemer, W.J., Ratamess, N.A., 2005. Hormonal responses and adaptations to resistance exercise and training. Sports Med. 35, 339–361.

Luhtanen, R., Crocker, J., 1992. A collective self-esteem scale: self-evaluation of one's social identity. Pers. Soc. Psychol. Bull. 18, 302–318.

Mazur, A., 1985. A biosocial model of status in face-to-face primate groups. Soc. Forces 64: 377–402. http://dx.doi.org/10.1016/j.sbspro.2013.06.508.

Mazur, A., Booth, A., Dabbs Jr., J., 1992. Testosterone and chess competition. Soc. Psychol. Q. 55:70–77. http://dx.doi.org/10.2307/2786687.

Mazur, A., Booth, A., 1998. Testosterone and dominance in men. Behav. Brain Sci. 21, 353–363.

Mazur, A., Lamb, T.A., 1980. Testosterone, status, and mood in human males. Horm. Behav. 14:236–246. http://dx.doi.org/10.1016/0018-506X(80)90032-X.

Mazur, A., Susman, E.J., Edelbrock, S., 1997. Sex difference in testosterone response to a video game contest. Evol. Hum. Behav. 18:317–326. http://dx.doi.org/10.1016/S1090-5138(97)00013-5.

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. J. Pers. Soc. Psychol. 94:1078–1093. http://dx.doi.org/10.1037/0022-3514.94.6.1078.

Mehta, P.H., Josephs, R.A., 2006. Testosterone change after losing predicts the decision to compete again. Horm. Behav. 50:684–692. http://dx.doi.org/10.1016/j.yhbeh.2006.07.001.

Mehta, P.H., Josephs, R.A., 2010. Testosterone and cortisol jointly regulate dominance: evidence for a dual-hormone hypothesis. Horm. Behav. 58:898–906. http://dx.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. Curr. Opin. Behav. Sci. 3:163–168. http://dx.doi.org/10.1016/j.cobeha.2015.04.008.

Mehta, P.H., Snyder, N.A., Knight, E.L., Lassetter, B., 2015. Close versus decisive victory moderates the effect of testosterone change on competitive decisions and task enjoyment. Adapt. Hum. Behav. Physiol. 1:291–311. http://dx.doi.org/10.1007/s40750-014-0014-0.

Mehta, P.H., Wuehrmann, E.V., Josephs, R.A., 2009. When are low testosterone levels advantageous? The moderating role of individual versus intergroup competition. Horm. Behav. 56:158–162. http://dx.doi.org/10.1016/j.yhbeh.2009.04.001.

Oliveira, T., Gouveia, M.J., Oliveira, R.F., 2009. Testosterone responsiveness to winning and losing experiences in female soccer players. Psychoneuroendocrinology 34: 1056–1064. http://dx.doi.org/10.1016/j.psyneuen.2009.02.006.

Oliveira, G.A., Uceda, S., Oliveira, T., Fernandes, A., Garcia-Marques, T., Oliveira, R.F., 2013. Threat perception and familiarity moderate the androgen response to competition in women. Front. Psychol. 4:1–8. http://dx.doi.org/10.3389/fpsyg.2013.00389.

Pfattheicher, S., Landhäuber, A., Keller, J., 2014. Individual differences in antisocial punishment in public goods situations: the interplay of cortisol with testosterone and dominance. J. Behav. Decis. Mak. 27, 340–348.

Popma, A., Vermeiren, R., Geluk, C.A.M.L., Rinne, T., Brink, W.V.D., Knol, D.L., Jansen, L.M.C., Engeland, H.V., Doreleijers, T.A.H., 2007. Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. Biol. Psychiatry 61, 405–411.

Rosenberg, M., 1965. Society and the Adolescent Self-Image. Princeton University Press, Princeton, NJ.

Salvador, A., 2005. Coping with competitive situations in humans. Neurosci. Biobehav. Rev. 29:195–205. http://dx.doi.org/10.1016/j.neubiorev.2004.07.004.

Salvador, A., Costa, R., 2009. Coping with competition: neuroendocrine responses and cognitive variables. Neurosci. Biobehav. Rev. 33:160–170. http://dx.doi.org/10.1016/j.neubiorev.2008.09.005.

Sapolsky, R.M., 2004. Social status and health in humans and other animals. Annu. Rev. Anthropol. 33:393–418. http://dx.doi.org/10.1146/annurev.anthro.33.070203.144000.

Schultheiss, O.C., Wirth, M.M., Torges, C.M., Pang, J.S., Villacorta, M.A., Welsh, K.M., 2005. Effects of implicit power motivation on men's and women's implicit learning and testosterone changes after social victory or defeat. J. Pers. Soc. Psychol. 88:174–188. http://dx.doi.org/10.1037/0022-3514.88.1.174.

Serrano, M.A., Salvador, A., González-Bono, E., Sanchís, C., Suay, F., 2000. Hormonal responses to competition. Psicothema 12, 440–444.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., Jacobs, G.A., 1983. Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press, Palo Alto, CA.

Stanton, S.J., Schultheiss, O.C., 2007. Basal and dynamic relationships between implicit power motivation and estradiol in women. Horm. Behav. 52:571–580. http://dx.doi.org/10.1016/j.yhbeh.2007.07.002.

Stockley, P., Bro-Jørgensen, J., 2011. Female competition and its evolutionary consequences in mammals. Biol. Rev. 86:341–366. http://dx.doi.org/10.1111/j.1469-185X.2010.00149.x.

Suay, F., Salvador, A., González-Bono, E., Sanchís, C., Martínez, M., Martínez-Sanchis, S., Simón, V.M., Montoro, J.B., 1999. Effects of competition and its outcome on serum testosterone, cortisol and prolactin. Psychoneuroendocrinology 24:551–566. http://dx.doi.org/10.1016/S0306-4530(99)00011-6.

Torrubia, R., Ávila, C., Moltó, J., Caseras, X., 2001. The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. Personal. Individ. Differ. 31:837–862. http://dx.doi.org/10.1016/S0191-8869(00)00183-5.

Trumble, B.C., Cummings, D., von Rueden, C., O'Connor, K.A., Smith, E.A., Gurven, M., Kaplan, H., 2012. Physical competition increases testosterone among Amazonian forager-horticulturalists: a test of the “challenge hypothesis”. Proc. Biol. Sci. 279: 2907–2912. http://dx.doi.org/10.1098/rspb.2012.0455.

Vallerand, R.J., 1983. The effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players. J. Sport Psychol. 5, 100–107.

van Anders, S.M., Watson, N.V., 2007. Effects of ability- and chance-determined competition outcome on testosterone. Physiol. Behav. 90:634–642. http://dx.doi.org/10.1016/j.physbeh.2006.11.017.

van der Meij, L., Buunk, A.P., Almela, M., Salvador, A., 2010. Testosterone responses to competition: the opponent's psychological state makes it challenging. Biol. Psychol. 84:330–335. http://dx.doi.org/10.1016/j.biopsycho.2010.03.017.

van Dijk, D., Kluger, A.N., 2011. Task type as a moderator of positive/negative feedback effects on motivation and performance: a regulatory focus perspective. J. Organ. Behav. 32, 1084–1105.

Watson, D., Clark, L.A., 1994. The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form. The University of Iowa, Ames, IA.

Webb, M.L., Wallace, J.P., Hamill, C., Hodgson, J.L., Mashaly, M.M., 1984. Serum testosterone concentration during two hours of moderate intensity treadmill running in trained men and women. Endocr. Res. 10, 27–38.

Welker, K.M., Lassetter, B., Brandes, C.M., Prasad, S., Koop, D.R., Mehta, P.H., 2016. A comparison of salivary testosterone measurement using immunosassays and tandem mass spectrometry. Psychoneuroendocrinology 71, 180–188.

West-Eberhard, M.J., 1979. Sexual selection, social competition, and evolution. Proc. Am. Philos. Soc. 123:222–234. http://dx.doi.org/10.1086/413215.

Wingfield, J.C., Hegner, R.E., Dufty Jr., A.M., Ball, G.F., 1990. The "challenge hypothesis": theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. Am. Nat. 126, 829–846.

Wirth, M.M., Welsh, K.M., Schultheiss, O.C., 2006. Salivary cortisol changes in humans after winning or losing a dominance contest depend on implicit power motivation. Horm. Behav. 49:346–352. http://dx.doi.org/10.1016/j.yhbeh.2005.08.013.

Zilioli, S., Watson, N.V., 2014. Testosterone across successive competitions: evidence for a 'winner effect' in humans? Psychoneuroendocrinology 47, 1–9.