Evolution and Human Behavior 37 (2016) 392–398

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

Does the facial width-to-height ratio map onto variability in men's testosterone concentrations?

Brian M. Bird [a], Valeska S. Cid Jofré [b], Shawn N. Geniole [c], Keith M. Welker [d], Samuele Zilioli [e], Dario Maestripieri [f], Steven Arnocky [g], Justin M. Carré [g,*]

[a] Department of Psychology, Laurentian University, Sudbury, ON, Canada
[b] Department of Ecology, University of Chile, Las Palmeras, Ñuñoa, Santiago, Chile
[c] Department of Psychology, Brock University, St. Catharines, ON, Canada
[d] Department of Psychology, University of Massachusetts Boston, Boston, MA, United States
[e] Department of Psychology, Wayne State University, Detroit, MI, United States
[f] Department of Comparative Human Development, University of Chicago, Chicago, IL, United States
[g] Department of Psychology, Nipissing University, North Bay, ON, Canada

ARTICLE INFO

Article history:
Initial receipt 20 July 2015
Final revision received 16 March 2016

Keywords:
Facial width to height ratio
fWHR
Testosterone
Competition
Neuroendocrine function

ABSTRACT

Variation in the facial width-to-height ratio (fWHR) maps onto a number of behavioral and psychological traits among men (e.g., aggression, unethical behavior, negotiation performance). Importantly, observer judgments of many of these traits also correlate strongly with the fWHR, suggesting that it may represent an honest cue to dominance and status. It has been speculated that the relationship between fWHR and these behavioral traits is due to pubertal testosterone concurrently shaping facial structure and traits linked to social dominance. Others, however, have provided some initial, although inconsistent, evidence that circulating testosterone levels in adulthood may underlie associations between the fWHR and behavioral displays. Here, we provide a more powerful test of the second model by examining the relationship between fWHR, baseline testosterone, and competition-induced testosterone reactivity, across seven diverse samples of men (total $N = 780$). We also report a further analysis including data published previously, for a total sample of 1041 men. Analysis of our individual samples, in addition to an internal meta-analysis, demonstrated no significant positive relationship between fWHR and baseline testosterone, or fWHR and three measures of competition-induced testosterone reactivity. We discuss potential reasons for previous discrepancies, and suggest avenues for future research.

1. Introduction

A growing body of evidence indicates that individual differences in facial morphology map onto a diverse range of behavioral and psychological traits, particularly among men (see Geniole, Denson, Dixson, Carre, & McCormick, 2015; Haselhuhn, Ormiston, & Wong, 2015, for meta-analyses). For instance, the facial width-to-height ratio (fWHR)—the distance of the bizygymatic width divided by the distance between the brow and upper lip—is positively correlated with measures of aggressive behavior (Carré & McCormick, 2008; Goetz et al., 2013; Lefevre et al., 2014; Welker, Goetz, Galicia, Liphardt, & Carré, 2014, but see Gómez-Valdés et al., 2013; Özener, 2012), psychopathic traits (Anderl et al., 2016; Geniole, Molnar, Carré, & McCormick, 2014), achievement drive (Lewis, Lefevre, & Bates, 2012), competitive success (baseball study of homeruns: Tsujimura & Banissy, 2013; formidability as a professional combatant: Trebická et al., 2015; Zilioli et al., 2014) unethical behavior (Geniole, Keyes, Carré, & McCormick, 2014; Haselhuhn & Wong, 2012), explicit prejudice (Hehman, Leitner, Deegan, & Gaertner, 2013), and negotiation performance (Haselhuhn, Wong, Ormiston, Inesi, & Galinsky, 2014). Notably, numerous studies also find that perceiver ratings of aggressiveness and dominance are highly correlated with the fWHR (e.g., Carré, McCormick, & Mondloch, 2009; Carré, Morrissey, Mondloch, & McCormick, 2010; Geniole, Molnar, et al., 2014; Short et al., 2012; see Geniole et al., 2015, for meta-analysis), suggesting that the fWHR may serve as a reliable cue to one's propensity for aggressive behavior.

It has been speculated that the link between facial structure and behavioral/psychological traits is due to the common influence of testosterone (T) on craniofacial growth and the expression of sexually dimorphic behaviors and traits (Carré & McCormick, 2008). Indeed, administration of T to males with delayed puberty modulates various indices of craniofacial growth (Verdonck, Gaethofs, Carels, & de Zegher, 1999). Other studies suggest positive associations between adult T concentrations and perceiver ratings of facial masculinity. In one study, composite images of men with high T (versus composites of those with low T) were rated as more masculine by observers (Penton-Voak & Chen, 2004). Similarly, men's T levels are positively correlated ($r =$

* Corresponding author. 100 College Dr., North Bay, Ontario, P1B 8L7.
E-mail addresses: justinca@nipissingu.ca, jcarre14@gmail.com (J.M. Carré).

http://dx.doi.org/10.1016/j.evolhumbehav.2016.03.004
1090-5138/

.34, $n = 38$) with female ratings of their facial masculinity (Roney, Hanson, Durante, & Maestripieri, 2006). However, more recent work has failed to find relationships between subjective ratings of facial masculinity and individual differences in baseline T concentrations (Peters, Simmons, & Rhodes, 2008).

Other research has examined links between objective measures of facial masculinity and individual differences in T concentrations, but results have been mixed (Apicella et al., 2008, 2011; Campbell et al., 2010; Lefevre, Lewis, Perrett, & Penke, 2013; Pound, Penton-Voak, & Surridge, 2009). For instance, Pound et al. found no relationship between facial masculinity (as measured through a global index of facial-masculinity) and baseline T concentrations in a small sample of young men ($r = 0.19$, $n = 47$). Instead, the authors reported a positive correlation between facial masculinity and T concentrations *after* watching a successful competitive interaction. In another study, Lefevre et al. found no relationship between baseline T and an objective measure of facial masculinity. In contrast, the authors found a small positive correlation between fWHR and baseline T concentrations ($r = .13$, $n = 188$) and between fWHR and T responses to a speed dating interaction (Lefevre et al., 2013). Moreover, Lefevre et al. (2013) found that fWHR was positively correlated with acute changes in T concentrations in response to a speed dating paradigm. Collectively, these studies provide mixed evidence for a link between facial structure and baseline T concentrations, and more consistent support for a relationship between facial structure and context-dependent fluctuations in T concentrations.

1.1. The present study

For the present study, we tested the previously reported positive associations between fWHR and T levels, using data from seven independent samples ($N = 780$). Further, we combined our seven samples with data that was previously published and publically available (Lefevre et al., 2013; http://www.sciencedirect.com/science/article/pii/S1090513813000275) to get the most robust test of a potential relationship ($N = 1041$). This study examined three main relationships of interest: (1) Is fWHR associated with baseline T? (2) Is fWHR associated with T responses to competition? (3) Does competition outcome (win vs. loss) moderate the relationship between fWHR and T reactivity? The latter question was motivated by evidence that demonstrated that changes in T concentrations during competition map onto variability in competitive motivation (Mehta & Josephs, 2006) and aggression (Carré, Putnam & McCormick, 2009) in losers, but not winners. To the extent that links between face structure and human behavior are mediated via neuroendocrine function, we wanted to examine whether the relationship between fWHR and T responses to competition would depend on the outcome of the competitive interaction.

2. Method

2.1. Participant samples

2.1.1. Sample 1

Photographs and T samples from 80 male participants between the ages of 18 and 33 ($M_{AGE} = 21.58$, $SD = 3.19$) were used from a previous study investigating testosterone responses to competition (see Norman, Moreau, Welker, & Carré, 2014, study 1, for full details). The majority of participants were Caucasian (86.1%), followed by Asian (5.1%), bi-racial/other (3.8%), First Nations/Aboriginal (2.5%), Black (1.3%), and Latin American (1.3%). Briefly, participants in the original study completed a video game task, which had been pre-programmed at a low level of difficulty, thus allowing participants to experience a string of victories. Saliva samples were collected pre (i.e., baseline) and post video game task.

2.1.2. Sample 2

Photographs and T samples from 114 male participants between the ages of 17 and 56 ($M_{AGE} = 21.78$, $SD = 5.68$) were examined from a previous study investigating testosterone responses to competition (see Carré, Campbell, Lozoya, Goetz, & Welker, 2013). T samples were available for 111 participants, and thus, the final sample size reflects this number. Participant ethnicities were classified as follows: Caucasian (52.6%), bi-racial/other (19.3%), Asian (18.4%), and Black (9.6%). In this study, participants were randomly and evenly assigned to a victory or defeat condition for a video game task. Saliva samples were collected at pre and post video game competition.

2.1.3. Sample 3

Photographs and T samples from 165 male participants between the ages of 18 and 34 ($M_{AGE} = 20.66$, $SD = 2.97$) were used from a larger protocol investigating hormones and competition (Welker & Carré, 2015). Full data for the variables of interest were available for 152 participants, and thus, this sample size was used for the present analysis. Ethnicities were diverse, with Caucasian (37.6%), Black (20%), Asian (18.2%), Middle Eastern (10.3%), bi-racial/other (8.5%), Latin American (4.8%), and First Nations/Aboriginal (0.6%) participants. Briefly, participants played an XBOX-360 video game randomly assigned to high difficulty (i.e., lose condition) or low difficulty (i.e., win condition), and gave a second saliva sample upon completion. Full details for this video game task are available in Carré et al. (2013).

2.1.4. Sample 4

Photographs and T samples available for 159 male participants between the ages of 24 and 35 ($M_{AGE} = 29.09$, $SD = 2.41$) were used from a previous study investigating the relationship between testosterone concentrations, risk aversion, and career choices among business school students (Sapienza, Zingales, & Maestripieri, 2009; see also Maestripieri, Baran, Sapienza, & Zingales, 2010). Participant ethnicities were varied, with Caucasian (43.4%), Native American (22%), Asian (13.8%), bi-racial/other (8.2%), Black (6.3%), and Latin American (6.3%). Participants in the original study had baseline T samples collected, and then engaged in a series of computerized decision-making tasks (Maestripieri et al., 2010; Sapienza et al., 2009). Since this study did not involve an experimental competition paradigm, only the baseline T concentrations were analyzed from this sample.

2.1.5. Sample 5

Photographs and T samples for 95 male participants between the ages of 18 and 30 ($M_{AGE} = 20.36$, $SD = 2.09$) were used from a study examining the relationship between competition, testosterone, and persistence in men (Welker & Carré, 2015). Ethnicities were classified as follows: Caucasian (46.3%), Black (21.1%), bi-racial/other (15.8%), Middle Eastern (6.3%), Asian (5.3%), Latin American (4.2%), and Native American (1.1%). In the original study, participants were randomly assigned to one of three conditions for a competitive number-tracing task: win against a confederate, lose against a confederate, or complete the task alone (control condition). Pre- and post-competition saliva samples were collected for hormonal assay.

2.1.6. Sample 6

Photographs and T samples were collected from 77 male participants aged 18 to 40 ($M_{AGE} = 21.84$, $SD = 3.56$). Participant ethnicities were classified as follows: Caucasian (75.3%), Middle Eastern (9.1%), Asian (9.1%) Black (3.9%), and Latin American. For this study, two saliva samples were collected in the afternoon between 2:00 pm and 4:00 pm across two consecutive days, and were averaged to create a mean T score.

2.1.7. Sample 7

Photographs and T samples were collected from 120 male participants between the ages of 18 and 35 ($M_{AGE} = 25.27$, $SD = 4.98$) that were part of a larger protocol examining the causal role of T on perception, cognition, and decision-making (Carré et al., unpublished). Briefly, participants reported to the lab and completed a battery of self-report