Hormones and Behavior 162 (2024) 105540

Contents lists available at ScienceDirect

Hormones and Behavior

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

Testosterone, gender identity and gender-stereotyped personality attributes☆

Kathleen V. Casto [a,*], Dale J. Cohen [b], Modupe Akinola [c], Pranjal H. Mehta [d]

[a] Kent State University, Department of Psychological Sciences, United States of America
[b] University of North Carolina Wilmington, Department of Psychology, United States of America
[c] Columbia University, Columbia Business School, United States of America
[d] University College London, Department of Experimental Psychology, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:
Testosterone
Sex differences
Gender identity
Masculinity
Femininity
Personality
Agency
Communality

ABSTRACT

Sex/gender differences in personality associated with gender stereotyped behavior are widely studied in psychology yet remain a subject of ongoing debate. Exposure to testosterone during developmental periods is considered to be a primary mediator of many sex/gender differences in behavior. Extensions of this research has led to both lay beliefs and initial research about individual differences in basal testosterone in adulthood relating to "masculine" personality. In this study, we explored the relationships between testosterone, gender identity, and gender stereotyped personality attributes in a sample of over 400 university students (65 % female assigned at birth). Participants provided ratings of their self-perceived masculinity and femininity, resulting in a continuous measure of gender identity, and a set of agentic and communal personality attributes. A saliva sample was also provided for assay of basal testosterone. Results showed no compelling evidence that basal testosterone correlates with gender-stereotyped personality attributes or explains the relationship between sex/gender identity and these attributes, across, within, or covarying out sex assigned at birth. Contributing to a more gender diverse approach to assessing sex/gender relationships with personality and testosterone, our continuous measure of self-perceived masculinity and femininity predicted additional variance in personality beyond binary sex and showed some preliminary but weak relationships with testosterone. Results from this study cast doubt on the activational testosterone-masculinity hypothesis for explaining sex differences in gender stereotyped traits and within-sex/gender variation in attributes associated with agency and communality.

1. Introduction

Sex or gender[1] differences in personality traits are commonly studied in psychological research yet the interpretation of results is the subject of ongoing debate (Del Giudice, 2023; Eagly and Revelle, 2022; Hyde et al., 2019; Kaiser et al., 2020). There appears to be general agreement that sex/gender differences exist, as evidenced by statistically significant effects across a wide array of personality facets and related behaviors and interests (Archer, 2019). However, disagreement lies in more subjective aspects of how data are analyzed and presented, how effect sizes are interpreted, and the strength of conclusions drawn (Del Giudice, 2022; Del Giudice et al., 2012; Maney, 2016). There also have been increased calls to include equal numbers of male and female animals and men and women in study designs and to test for sex/gender differences, including formal directives from research funding agencies (e.g., Woitowich et al., 2020). Doing so necessitates the identification of best practices for testing sex/gender differences as well as innovative methods for more gender-diverse approaches to assessing sex/gender as a variable in research designs (Garcia-Sifuentes and Maney, 2021; Maney and Rich-Edwards, 2023).

A separate but related debate also exists over what factors are most influential in producing sex/gender differences in personality (Eagly

☆ This paper is part of the Virtual Special Issue Sex/gender Diversity and Behavioral Neuroendocrinology in the 21st Century.

* Corresponding author.
E-mail address: kcasto6@kent.edu (K.V. Casto).

[1] Although a subject of debate, the term "sex" is often used when referring to instances when anatomical, histological, genetic, or hormonal are the focus of classification; "gender" is used when referring to socio-cultural systems which describe the ways sex interacts with culture, psychology, and history to produce gendered self- identity and performative acts (Muehlenhard and Peterson, 2011; Hyde et al., 2019). When the distinction is not clear, as is often the case, we use "sex/gender" or the specific sex/gender related variable that was measured in this study.

https://doi.org/10.1016/j.yhbeh.2024.105540
Received 30 October 2023; Received in revised form 29 March 2024; Accepted 29 March 2024
Available online 22 April 2024

and Wood, 2013; Schmitt et al., 2017). Personality is a complex array of traits and behavioral tendencies brought about by an amalgam of genetic and environmental factors and changeable across the social and developmental context (Briley and Tucker-Drob, 2014; Hopwood et al., 2011; Vukasovic and Bratko, 2015). However, sex/gender-linked differences in personality are often examined with a more singular focus on evolutionary processes or patterns of developmental hormone exposure (Collaer and Hines, 1995; Hines, 2002; Puts, 2016; Shirazi et al., 2022). Indeed, hormones are viewed as the primary explanatory agent of sex-differentiated social behavior (Williams et al., 2023), e.g., "research is thus consistent with the claim that hormonal mechanisms mediate the sexes' contingent responding to their social relationships" (p. 86, Wood and Eagly, 2012). Yet, gendered self-expression is influenced by many factors including social experiences which are deeply entangled with hormonal and genetic factors; and, the relative importance of hormonal mediation remains in question.

Prominent research in the field of behavioral neuroendocrinology has provided the basis for emphasizing the role of testosterone, specifically, in organizing sex differences in behavior in rodents and primates, including humans (for review, Berenbaum and Beltz, 2011; McCarthy et al., 2009; Morris et al., 2004; Swift-Gallant et al., 2023; Wallen, 2005). In fact, the field was founded in part by seminal studies demonstrating testosterone's effects on sex-typical sexual behavior in rodents (Phoenix et al., 1959). Evidence for the effects of early androgen exposure on human behavior comes largely from the study of individuals with a genetic condition (Congenital Adrenal Hyperplasia, CAH) in which chromosomal females are exposed to relatively high levels of androgens compared to typically-developed females (for review, Berenbaum and Beltz, 2011, 2021). In these studies, CAH women show "masculinized" behavior including aspects of childhood play, preferred activities, and interests (e.g., Berenbaum, 1999; Berenbaum and Hines, 1992; Nordenström et al., 2002; Servin et al., 2003). In other studies not involving people with CAH, relatively higher blood levels of testosterone in pregnant mothers has been shown to predict more masculine-typical gender role behavior of the children (Auyeung et al., 2009; Hines et al., 2002). More recent research has found that men with another condition that leads to deficient androgen exposure in early development recalled less masculine conforming childhood behaviors than unaffected men (Shirazi et al., 2022).

However, there is a critical distinction between testosterone's "*organizational*" effects on behavior which result from early exposure during development and the "*activational*" effects on behavior, a broader class of effects having to do with adulthood exposure patterns and responsivity to environmental and situational demands (for review, Arnold, 2009). Conflating the two can lead to misunderstandings and biased predictions about how individual differences in testosterone in adulthood relate to sex differences in behavior and personality (Hyde et al., 2019; van Anders, 2013). One primary example is the notion that adults with relatively higher levels of testosterone have more "masculine" traits, e.g., competitive, agentic, and assertive, and lower "feminine" traits, e.g., communal, devoted to others, and emotional (e.g., Jordan-Young and Karkazis, 2019). Specifically, women with relatively high basal testosterone compared to typical women are thought to be more stereotypically male-like, while men with relatively low basal testosterone are thought to be more stereotypically female-like, compared to most men or women, respectively – herein termed the *testosterone-masculinity hypothesis*.

Some early research has been conducted to test these predictions. For example, Baucom et al. (1985) reported that women with higher testosterone levels self-identified as more "self-directed, action-oriented, and resourceful" while those with relatively lower levels "viewed themselves as conventional, socialized individuals, possessing a caring attitude coupled with an anxious and dejected mood." More recent research has continued this trend. A study of over 2000 men with clinical sexual dysfunction found that testosterone levels in blood were positively related ($r = 0.08$) to a histrionic/hysterical traits subscale of a standard personality inventory, a finding that contradicts gender stereotypes about testosterone (Bandini et al., 2009). However, the authors maintained a strong and causal view of testosterone's influence on personality by concluding in the title of the report that "hysterical traits are not from the uterus but from the testis."

The testosterone-masculinity hypothesis is the foundation of closely related areas of study where testosterone variation has been proposed as an explanation for why some, mostly men, are more likely to commit violent crimes than others (Armstrong et al., 2022; Banks and Dabbs, 1996; Dabbs and Morris, 1990; Ehrenkranz et al., 1974) and why some individuals prefer same-sex sexual partners (for review, Cunningham and Benítez, 2024). One line of research suggests that testosterone levels play a role in promoting "honor cultures" and violence among "uneducated" Black men (Mazur, 2016). Beyond its potentially harmful social implications, the testosterone-masculinity hypothesis has real-world consequences for health and wellbeing. In both explicit and tacit acknowledgement of testosterone's suspected effects on personality, testosterone replacement therapy for "low T" men has become a multi-billion-dollar industry. In the introduction to the popular book "The Virility Factor: Masculinity Through Testosterone, the Male Sex Hormone" (Bahr, 1992, p. xi), medical endocrinologist and early adopter of the practice of testosterone therapy Hubert Kupperman wrote,

> "…increased testosterone levels above the norm occur in the dominant male who is the leader of the pack. When the male hormone is deficient…there is a tendency for these males to accept a passive role…When testosterone levels are inadequate, physical debilitation associated with poor muscular development may occur, and *changes in the personality may be seen*. It is for this reason that wise and early implementation of appropriate therapy would be helpful in preventing some disastrous psychological aberrations from taking place in the male."

Indeed, advertisements for testosterone treatment prey on this persistent and insidious meme that testosterone imbues masculinity, that a man is not one without it, "…it is anyone's guess, whether he is slave or master of the testosterone that makes him a man… And man is in all his essence the product of his hormones. (p. 42, Bahr)." Yet, the psychological and health risks and benefits of testosterone therapy remain an active area of debate (Rodrigues dos Santos and Bhasin, 2021; Snyder et al., 2018). Misunderstandings about individual differences in adulthood testosterone levels are also at the core of recent discriminatory practices in sport in which some regulatory bodies have used testosterone levels as a singular criterion for exclusion in the female classification (Casto and Carre, 2023).

Although there is limited prior evidence in support of the testosterone-masculinity hypothesis, there is a more robust literature on the relationship between individual differences in basal testosterone and personality more broadly (e.g., Sellers et al., 2007). One personality trait that has been particularly well studied in this area is extroversion, which may be important for a social-approach orientation underlying status-seeking and dominance (e.g., behaviors thought to be influenced by testosterone). Although some studies have shown a positive correlation between extroversion and testosterone levels (Alvergne et al., 2010; Crewther et al., 2020), these correlations appear relatively small ($r \sim 0.10$ at best across different facets of the trait; Smeets-Janssen et al., 2015). Further, a recent meta-analysis of studies on this association found no overall support (Sundin et al., 2021). Other meta-analyses and analytic reviews on testosterone and self-reported traits associated with social dominance and status motivation have similarly concluded that relationships are weak or non-significant (Casto et al., 2023; Grebe et al., 2019a). Yet there remains a need to directly test predictions about testosterone's relationship to stereotypically masculine and feminine traits, an underlying tenet of the testosterone-masculinity hypothesis, in an adequately powered sample of both men and women.

An additional and broader gap in prior literature is the general lack of studies that test the strength of testosterone as a mediator *in*

comparison to the potential role of psychosocial mediators. Such comparison allows for the consideration of alternative hypotheses and counter-theory to more strictly hormonal approaches. Some studies of testosterone's role in explaining sex effects on spatial cognition have provided examples for testing the mediating and moderating roles of psychosocial factors, including indices of masculinity and femininity (Constantinescu et al., 2018; Hausmann et al., 2009; Pletzer et al., 2019; Puts et al., 2010). These studies provide initial evidence that both testosterone and psychosocial factors like gender-role expectations or perceptions of masculinity and femininity may influence sex differences in performance on spatial cognition tasks. Other studies have tested the strength of self-perceived masculinity as a mediator of sex effects on spatial cognition (i.e., mental rotations), consistent with the sex-role mediation hypothesis (Kelly and Beltz, 2022). Yet, to the best of our knowledge, no research to date has directly examined competing models for the interplay between testosterone levels, sex/gender, masculinity/femininity, and gendered personality.

The goal of the present study is to explore the relationships between basal testosterone levels, self-perceptions of masculinity and femininity, and a set of agentic and communal-based attributes traditionally ascribed to gender categories. To account for variability in sex/gender identity and to heed calls for more gender non-binary approaches in this area of research , we employ a continuous measure of sex/gender which is based on self-perceptions of masculine and feminine body and personality features. We examine basal testosterone's associations to these gender identifications and the gender-stereotyped personality attributes overall and separately within sex/gender category. Next, we test and compare the predictive strength of three models: 1) Basal testosterone as a mediator of the effect of our continuous measure of gender identity on gender-stereotyped (GS) personality, 2) Gender identity as a mediator of the effect of testosterone on GS personality, and 3) The interaction term of gender identity and testosterone in predicting GS personality. Initial analyses without controlling for categorical sex assigned at birth test the magnitude with which testosterone relates to gender identity and personality overall and, if so, whether it statistically explains sex/gender differences in personality. Then, we conduct these analyses while controlling for categorical sex/gender, which enables us to determine whether testosterone correlates with personality beyond the sex/gender binary and mediates the effect of within-sex variation in gender identity on gendered personality. Overall, our approach is correlational and exploratory and therefore, cannot be interpreted as indicating causality about how testosterone and gender identification relate to gendered behavior, but the findings can inform future study designs that aim to clarify the causal pathways.

2. Method

2.1. Participants

A sample of 434 undergraduates (mean age of 20 years, SD = 4.0; 65 % assigned female at birth) from the University of Oregon participated in the study in the spring of 2017 through the spring of 2018. Sample size was determined by the goal to recruit as many participants as possible within three semesters. A power sensitivity analysis (R package pwr2ppl) revealed 80 % power to detect an effect size of $r = 0.17$–0.18 for standard mediation with three variables.

Participants were asked two questions about categorical sex and gender. The first question stated, "What sex category were you assigned at birth?" followed by the answer choices "Male," "Female," "Intersex," and "Other." From the sample, 152 identified as assigned male at birth (AMAB) and 282 identified as assigned female at birth (AFAB). The second question stated, "What is your current gender identity?" followed by the answer choices "Woman," "Man," "Transgender woman," "Transgender man," "Genderqueer," and "None of the above." From the sample, 275 identified as women, 151 as men, 1 as transgender woman, 2 as transgender man, 4 as genderqueer, and 1 as gender not listed. Participants were also asked to respond yes or no to the question "Are you taking any hormonal supplements?" If they responded "yes," then they were asked to subsequently write in a text box what they were taking. One transgender man and woman were on gender affirming hormonal therapy and were excluded from analyses due to the inadequate sample size for testing effects of exogenous steroids.

Ethnicity was predominantly European/European-American (61 %) with relatively moderate subsets of Hispanic/Latino (14 %) and Asian/Asian-American (14 %). Participants were recruited from the psychology department subject pool composed of undergraduates enrolled in introductory psychology and linguistics courses, each of which has a research participation option as a condition for the satisfactory completion of the course. The study was approved by the university's Research Compliances Services.

2.2. Measures

2.2.1. Continuous measure of gender identity

An aim of this study was to explore methods of operationalizing and analyzing sex/gender identity beyond the binary framework. To do this, we asked participants to rate their self-perceived masculinity and femininity. Further, participants made these ratings for two aspects of gendered self-perception, the physical body and the personality (a similar approach was taken by Pletzer et al., 2015, 2019 with an even greater number of sub-categories). Body was included in addition to personality due to evidence that gender identification is tightly connected with body perception (Clausen et al., 2021; Tacikowski et al., 2020). Participants were asked to "Move each of the sliders to rate the degree with which you view your BODY as masculine and feminine on a continuum" and "Move the slider to rate the degree with which you view your PERSONALITY as masculine and feminine on a continuum." Below each instruction were the words Masculine and Feminine and a corresponding slider scale ranging from 0 to 100 with three text labels "Less true of me," "Moderately," and "More true of me". Thus, four total ratings were made.

As shown in Fig. 1, ratings of body femininity and personality femininity were highly positively correlated ($r = 0.86$), as were body and personality masculinity ($r = 0.79$). The body and personality item ratings were summed as feminine gender identity (GenFem) and masculine gender identity (GenMas). Cronbach's alpha for GenFem = 0.92 and for GenMas = 0.88. The scale items are face-valid in that they are direct and appear to the respondent in the same way that the responses are operationalized. As an additional check for the purpose of providing validity evidence, we tested the correlation between the GenFem and GenMas scores with a validated measure of masculine and feminine "gender typicality" (Egan and Perry, 2001; Patterson, 2012). Correlations with the masculine and feminine typicality were significantly and positively related to GenMas ($r = 0.48, p$ .001) and GenFem ($r = 0.33, p$ .001). The response distributions for GenFem and GenMas by categorical sex assigned at birth are shown in Fig. 2. As evident in both Figs. 1 and 2, there is some overlap between men and women on these aspects of gender identity as well as variability along the continuum that are obfuscated by dichotomizing sex/gender. Nonetheless, categorical sex assigned at birth was highly predictive of self-perceptions of gender (GenMas = 0.86; GenFem = 0.93; both p .001) and thus, the resulting constructs are conflated with sex category assigned at birth.

2.2.2. Personal Attributes Questionnaire (PAQ)

The PAQ (Helmreich et al., 1981; Spence and Helmreich, 1978) is a self-report measure designed to assess the degree to which a person identifies as having stereotypically masculine traits (agency and instrumentality) and feminine traits (communality and expressiveness). Respondents indicated their agreement that a list of 16 adjectives (Table 1) were self-descriptive ("the degree with which each trait describes you") using a slider scale from 0 to 9 with 7 labels ranging from

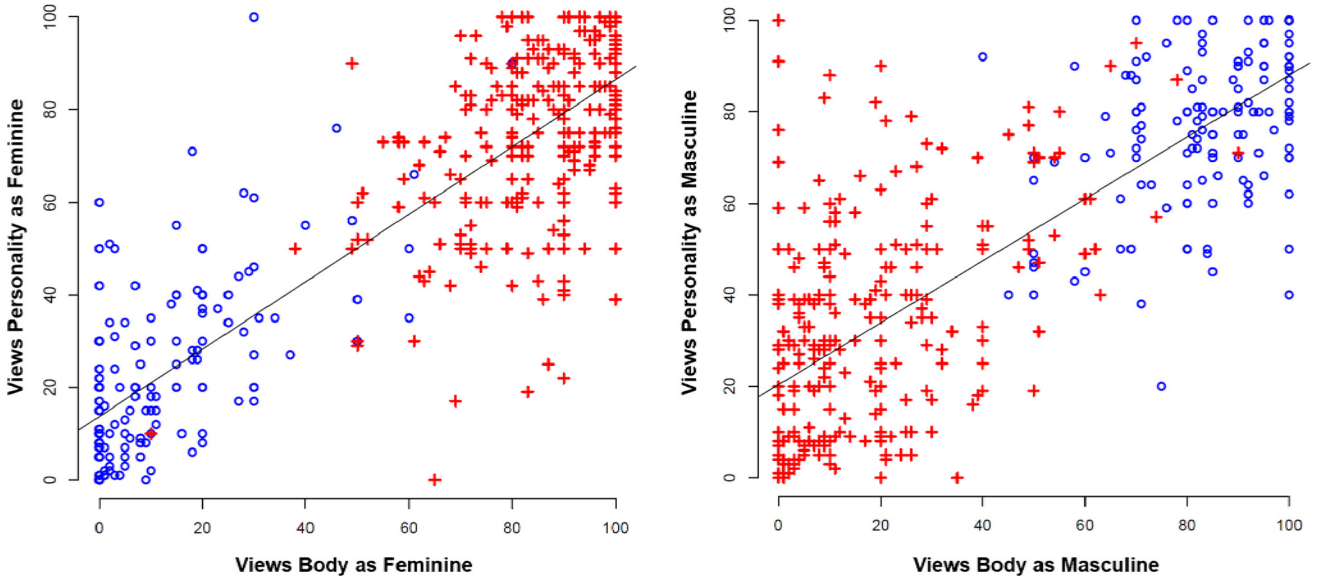

Fig. 1. Scatterplots of the relationships between the degree with which female- (cross) and male- (circle) assigned at birth participants viewed the femininity of their body and personality (left figure) and masculinity of their body and personality (right figure).

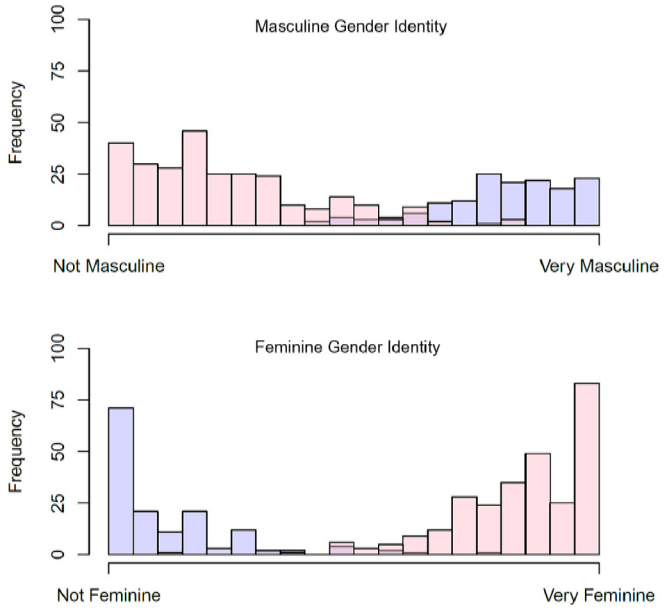

Fig. 2. Histogram for masculine gender identity (Gen mas = body + personality) and feminine gender identity (Gen fem = body + personality) by sex assigned at birth. Red bars = AFAB; Blue bars = AMAB.

"Strongly Disagree" to "Strongly Agree". This scale has been well-validated, although the factor structure is less stable in cultures outside of US populations (Hill et al., 2000; McCreary and Steinberg, 1992). In our sample, Cronbach's alpha for the full PAQ-feminine scale (α = 0.81) and PAQ-masculine scale (α = 0.75) were acceptable. An individual's mean self-rating on the PAQ-masculine and PAQ-feminine subscales was calculated. Due to a tradition for a two-dimensional structure for these types of scales (Choi et al., 2007; Wood and Eagly, 2015; but see, Gaa et al., 1979), we conducted main analyses with PAQ-mas and PAQ-fem as separate outcomes variables.

Research has shown that stereotypically masculine and feminine traits have changed alongside changing gender role expectations in the decades since the development of the PAQ and similar inventories (Bhatia and Bhatia, 2021; Conway and Vartanian, 2000; Donnelly and Twenge, 2017; Hentschel et al., 2019). Thus, a sex/gender difference on specific attributes on the PAQ are not expected for all attributes. We conducted an assessment of the sex/gender differences on each attribute and the aggregate of the masculine and feminine items within our sample. After correcting for multiple tests (Bonferroni alpha = 0.006), significant differences between men and women, with moderate effect sizes but high percent overlap, emerged for four of the eight attributes in each subscale (shaded in Table 1).

2.3. Procedure and salivary assays

Participants were instructed not to eat, exercise, smoke, or consume caffeinated beverages or food within the hour prior to arriving at the laboratory for the study. All sessions were scheduled between 1 and 4 PM to control for circadian effects on basal testosterone levels. Upon arrival at the laboratory, participants were given an overview of the study and completed the consent form. Then, they were placed in a private room to complete surveys for 20 min. The survey included the demographic questions (i.e., assessment of participants' sex assigned at birth and current gender identity), the PAQ, the gendered-self scales and other inventories (assessed for research unrelated to the present study). Afterwards, participants were retrieved from their room and instructed to rinse their mouth with water. They were then given a plastic vial and a paper towel, and detailed instructions on the process for producing passive drool. Approximately 1–3 ml of saliva was collected for each sample via passive drool into 5 ml plastic vials. Participants were given up to7 min to produce the saliva sample.

Saliva samples were stored at 20 °C initially and then transferred to a 80 °C freezer for 1–6 months before being shipped on dry ice to Dresden, Germany. Samples were assayed for testosterone by Dresden Lab Service using IBL chemiluminescence immunoassay kits with 20 % of samples tested in duplicate. This research was a part of a larger study (e.g., Casto et al., 2020) that also included measurement of cortisol and DHEA-S levels. Only testosterone levels are included for analysis in the present study. Testosterone CV% for low and high controls were 4.86 % and 4.29 %, respectively and intra-assay CV% was 4.5 %. The sample size for full analyses was 419–421 due to saliva sample loss/degradation (a few samples were destroyed in shipment) and incomplete survey responses.

Table 1
Sex/gender differences and similarities in PAQ masculine and feminine items.

Personal attribute	AMAB *M(SD)* *N* = 151	AFAB *M(SD)* *N* = 281	Cohen's D*	% overlap*	*t(df)*	*p*	*CI*	
Feminine (communal and expressive)								
Emotional	4.95 (2.24)	6.09 (2.02)	0.53	78%	5.22 (281)	.001	1.57	0.71
Devoted to others	5.91 (1.95)	6.62 (1.73)	0.38	84%	3.70 (277)	.001	1.07	0.33
Warm	6.07 (1.97)	6.66 (1.86)	0.31	87%	3.04 (292)	.003	0.98	0.21
Kind	7.12 (1.55)	7.57 (1.31)	0.32	86%	3.06 (266)	.002	0.75	0.16
Understanding	7.26 (1.64)	7.62 (1.39)	0.23	88%	2.26 (267)	.025	0.66	0.05
Helpful	6.86 (1.59)	7.18 (1.52)	0.20	92%	2.01 (297)	.046	0.63	0.01
Aware of feelings	6.85 (2.08)	7.10 (1.82)	0.12	92%	1.20 (275)	.23	0.64	0.15
Gentle	6.29 (1.92)	6.13 (1.79)	0.09	95%	0.84 (290)	.399	0.21	0.53
Total scale	6.41 (1.28)	6.87 (1.04)	0.39	82%	3.76 (259)	.001	0.69	0.22
Masculine (instrumental and agentic)								
Competitive	5.96 (2.26)	4.90 (2.49)	0.45	82%	4.48 (333)	.001	0.59	1.53
Decisive	5.19 (2.19)	4.53 (2.14)	0.31	88%	3.03 (301)	.003	0.23	1.10
Feels superior	3.68 (2.46)	2.98 (2.19)	0.30	88%	2.93 (278)	.004	0.23	1.17
Active	6.53 (1.95)	5.89 (2.05)	0.33	88%	3.18 (321)	.002	0.24	1.03
Stands up under pressure	5.91 (2.11)	5.66 (2.19)	0.11	96%	1.13 (317)	.257	0.18	0.67
Self-confident	5.88 (2.47)	5.67 (2.31)	0.09	96%	0.86 (290)	.393	0.27	0.69
Never gives up	6.24 (2.27)	6.13 (2.03)	0.05	94%	0.48 (280)	.629	0.33	0.54
Independent	6.72 (1.83)	7.12 (1.69)	0.23	90%	2.26 (286)	.025	0.76	0.05
Total scale	5.76 (1.38)	5.36 (1.26)	0.30	88%	2.98 (284)	.003	0.14	0.67

Note. AMAB = self-reported as assigned male at birth; AFAB = self-reported as assigned female at birth. *Calculated via sexdifference.org.

2.4. Statistical approach

Testosterone concentrations were not normally distributed (positively skewed), both within and between sex. Raw values were log transformed to reduce the skew. Nine participants' testosterone levels were identified as outliers for being 3 SDs above the mean for their sex assigned at birth category; their testosterone value was winsorized to the 3 SD mark. Salivary androgens measured via enzyme and chemiluminescent immunoassay systematically overestimate women's levels and artificially reduce the sex differences in testosterone levels measured in blood (Chafkin et al., 2022; Welker et al., 2016). The log transformations reduced the otherwise robust sex difference in testosterone levels even further. Because the purpose of the present study was to assess the correlates of basal testosterone in a mixed-sex sample, as a robustness test to the results despite these analytic choices, we also repeated analyses with raw testosterone concentrations. We reported whether these robustness tests were consistent with the results of the primary analyses in the results and provide results for raw testosterone in the Supplemental file.

Standard *t*-tests and Pearsons' correlations were used to calculate descriptive statistics. We used the PROCESS macro in R to compute the effect size for mediation (Model 4) and to compute the bootstrapped confidence intervals of the indirect effect (Hayes, 2020; Hayes and Scharkow, 2013; Tibbe and Montoya, 2022). For our Model 1, gender identity was the predictor (X), testosterone was the mediator (M), and personality attributes were the outcome (Y). In our Model 2, testosterone was the predictor (X), gender identity was the mediator (M), and personality attributes were the outcome (Y). The variables and paths for analysis for Models 1 and 2 are shown in Fig. 3. In our Model 3, linear regressions that predicted personality attributes were conducted where gender identity and testosterone were entered together in step 1 and their interaction term was entered in Step 2. Due to the two-factor structure for our measures of gender identity and personality attributes, each model was run separately for the corresponding masculine and feminine identity and attribute scale. To test the effects in the above models after controlling for categorical sex assigned at birth, we also provided results for the models with sex included as a covariate. Given the potential for multicollinearity among the predictors and covariate, we also tested the variance inflation factor (VIF) for all main analyses. Secondary analyses repeated the above analyses separated by sex assigned at birth to assess the effects of within-sex individual differences in gender identity and testosterone on masculine and feminine personality attributes. Finally, for analyses with assigned-female-at-birth participants, secondary analyses were repeated with hormonal contraceptive (HC) use (yes/no) as a covariate, due to known effects of HCs reducing testosterone levels (e.g., Zimmerman et al., 2014).

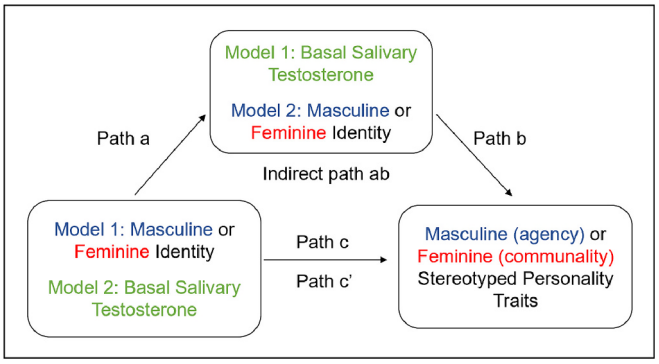

Fig. 3. Variables and paths for mediation models 1 and 2.

Due to the high number of tests, we report whether the *p*-value was 0.01 or 0.001, a more conservative alpha, and focus interpretations on the path weights, effect sizes, and robustness of effects (consistent with recommendations by Cumming, 2014). It is well known that conventional mediation analyses with unmanipulated mediators have a high likelihood of bias (e.g., Bullock et al., 2010). This bias will inflate the influence of the mediator when the mediator is correlated with an unobserved variable that influences both the mediator and the DV or simply the DV alone (Bullock and Green, 2021). Because our mediators are unmanipulated, the size of the mediation effects should be interpreted with caution. Further, there are important limitations to the interpretations of exploratory model comparisons when using the statistical approaches employed in this study (for review, Rohrer et al., 2022). Thus, results should not be interpreted as evidence of causal inference.

All statistical analyses were conducted in R (version 4.2.1). All data, materials, and analysis code are publicly posted on the open science framework (https://osf.io/bny4j/).

3. Results

3.1. Descriptives and correlations

Descriptive statistics and correlations for the main study variables are shown in Table 2, overall and within sex. A more feminine gender identity was related to higher scores on feminine-stereotyped personality and lower masculine-stereotyped personality, but effects were small to moderate in size. The same relationships and effect sizes, in reverse, were observed for masculine gender identity. Feminine and masculine gender identities were strongly anti-correlated, while masculine and feminine personality attributes showed a small positive correlation.

In a mixed-sex sample, testosterone was not correlated with PAQ masculine traits ($R^2 = 0.01$; not significant) and only weakly inversely related to PAQ feminine traits ($R^2 = 0.02$; significant). These small trends were eliminated when controlling for sex assigned at birth. There was a moderate negative correlation between testosterone and feminine gender identity ($R^2 = 0.52$) and a moderate positive correlation between testosterone and masculine gender identity ($R^2 = 0.46$). When controlling for sex assigned at birth, correlations between testosterone and the gender identity measures were in the same direction, but diminished ($R^2 = 0.01$–0.02). These weak trends when controlling for sex were reduced further and were no longer significant when using raw testosterone in place of the log-transformed values.

In AFAB individuals only, those with higher testosterone reported a slightly more masculine ($R^2 = 0.02$) and less feminine gender identity ($R^2 = 0.01$), but effects were not significant (results were the same with and without controlling for HC use). A similar negative trend was seen for testosterone and feminine gender identity in AMAB individuals, but the correlation between testosterone and masculine gender identity was close to zero. Finally, there was an unexpected, yet not significant trend for AMAB participants with higher testosterone to score higher on feminine-stereotyped personality attributes. Overall, and especially when accounting for sex assigned at birth, there was no compelling evidence that individual variability in basal testosterone[2] was meaningfully related to self-perceived masculinity or femininity or, masculine or feminine-stereotyped personality traits.

3.2. Model comparisons

Three models were tested to explore the role of testosterone and gender identity in predicting gender-stereotyped personality traits (PAQ fem and PAQ mas): Model 1 tests whether sex differences in basal testosterone mediate the effects of gender identity on personality, Model 2 tests whether gender identity mediates the effect of sex/gender differences in testosterone on personality, and Model 3 tests whether the interaction term of gender identity and testosterone predict personality. Repeating these analyses with categorical sex/gender as a covariate effectively tests how each model predicts personality beyond or independent of binary sex/gender. Results for the main analyses with the full sample of all participants are shown in Table 3. Results for the sex-covariates analyses should be interpreted with caution due to moderate to high variance inflation factor, i.e., effects may be inflated due to multicollinearity.

As identified in the correlations in Table 2, in the combined sample of men and women, the "a paths" in Models 1 and 2 were significant; higher basal testosterone levels predicted a more feminine and less masculine gendered identity in the full sample. Including sex as a covariate reduced the magnitude of these relationships and they were no longer significant. In Model 1, there was a positive and moderate direct effect of masculine and feminine identity on their respective masculine or feminine-stereotyped personality attributes (with and without controlling for sex). There was no supporting evidence of mediation by testosterone. The results for Model 1 were the same when raw testosterone levels were substituted for log-transformed testosterone and when categorical sex was substituted for our continuous measures of gender identity (results in the Supplement).

In Model 2, the total effect for testosterone on masculine attributes was only weakly positive and non-significant, but masculine identity mediated this small effect. The total effect for testosterone on feminine attributes was weakly negative and significant, but feminine identity was also a mediator of this effect. In both instances, there was no direct effect of testosterone on masculine or feminine attributes, after accounting for variance due to the corresponding gender identity. When controlling for sex, there was no compelling evidence for robust direct or indirect effects, but our continuous measure of masculine and feminine gender identity remained a moderate predictor of their corresponding masculine and feminine personality traits above and beyond categorical sex assigned at birth. The results for Model 2 were the same when raw testosterone levels were substituted for log-transformed testosterone and when categorical sex was substituted for our continuous measures of gender identity (results in the Supplement).

Finally, Model 3 showed that with and without controlling for sex, there was no compelling evidence that masculine gender identity and testosterone interact to predict masculine attributes. There was a small moderation effect (1–2 % additional variance explained) such that feminine gender identity and testosterone interact to predict feminine attributes. To interpret this interaction effect, we conducted simple slopes analyses. Among participants with low feminine gender identity (1 SD) and at the mean on feminine identity, testosterone was positively linked to feminine personality (1 SD: $b = 0.46$, $SE = 0.13$, $t = 3.63$, $p = .003$, $CI = 0.21$–0.72; mean: $b = 0.22$ $SE = 0.08$, $t = 2.71$, $p = .007$, $CI = 0.06$–0.37). Among participants with high feminine gender identity (+1 SD), testosterone was not linked to feminine personality ($b = 0.03$, $SE = 0.09$, $t = 0.37$, $p = .712$, $CI = 0.21$–0.15). However, the interaction effect was in the same direction, but dampened and no longer significant when controlling for categorical sex and when raw testosterone levels were substituted for log-transformed testosterone.

Secondary analyses of the model comparisons within each sex assigned at birth category (Tables 4 and 5) revealed that the relationships between testosterone and gender identity (a paths), were no longer evident. Further, there was no strong evidence for direct or indirect effects of testosterone on gender-stereotyped personality attributes and, among AFAB participants, results were essentially the same when including HC use as a covariate (see Supplemental analyses). The effect of gender identity on the corresponding personality attribute scales showed some additional nuance in the within-sex analyses. In AFAB participants, there was a moderate direct and positive effect of feminine identity on feminine attributes, but no significant effect of masculine gender identity on masculine attributes. In the reverse, AMAB participants showed a moderate direct and positive effect of masculine identity on masculine-stereotyped personality, but no significant effect of feminine identity on feminine-attributes.

4. Discussion

The purpose of the present study was to explore the relationships between sex/gender identity, individual variability in basal testosterone, and gender-stereotyped personality traits. A significant sex/gender difference in these so called "feminine" and "masculine" traits was only found for four of the eight attributes in each category. The effect size for these differences was moderate (Cohen's d = 0.30–0.53), but still showed a 78–88 % overlap in the distributions of men and

[2] Basal cortisol levels were also collected with this dataset. Due to prior evidence that cortisol moderates testosterone's relationship with status behavior (Knight et al., 2020; Mehta and Josephs, 2010), we include results for cortisol and the interaction of cortisol and testosterone in the Supplement. Neither cortisol nor its interaction with testosterone predicted the masculine or feminine personality attributes or gender identity.

Table 2
Sample sizes, ranges, means, standard deviations, and correlations among the main study variables. Partial correlations controlling for sex assigned at birth are included in parentheses below to each simple correlation in the full sample matrix.

	N	Min–max	M (SD)	T	PAQ mas	PAQ fem	Gen Mas
				Full sample			
Testosterone (pg/ml)	421	1.27–237.00	50.36 (45.60)	–			
PAQ Mas	432	1.88–8.38	5.50 (1.31)	0.08 (0.05)	–		
PAQ Fem	432	1.12–9.00	6.71 (1.15)	0.15* (0.01)	0.14* (0.17**)	–	
Gen Mas	429	0–200	88.44 (63.82)	0.68** (0.13*)	0.19** (0.12)	0.21** (0.10)	–
Gen Fem	432	0–200	119.31 (75.47)	0.72** (0.11)	0.12 (0.05)	0.23** (0.15*)	0.88** (0.43**)
				AFAB			
Testosterone (pg/ml)	273	1.27–91.00	24.52 (19.17)	–			
PAQ Mas	281	1.87–8.25	5.36 (1.26)	0.07	–		
PAQ Fem	281	3.63–9.00	6.87 (1.04)	0.04	0.22**	–	
Gen Mas	278	0–165	48.17 (35.66)	0.14	0.01	0.19**	–
Gen Fem	281	20–200	170.45 (27.92)	0.12	0.15	0.23**	0.39**
				AMAB			
Testosterone (pg/ml)	148	29.27–237	98.04 (41.63)	–			
PAQ Mas	151	2.0–8.38	5.76 (1.38)	0.01	–		
PAQ Fem	151	1.13–8.75	6.41 (1.28)	0.14	0.11	–	
Gen Mas	151	85–200	162.60 (27.07)	0.01	0.36**	0.07	–
Gen Fem	151	0–160	24.15 (29.93)	0.10	0.13	0.02	0.55**

Note. * $p \leq .01$; ** $p \leq .001$. PAQ Diff = PAQ feminine – PAQ masculine. AMAB = self-reported as assigned male at birth; AFAB = self-reported as assigned female at birth. For all correlations, the results are similar in magnitude and direction when raw testosterone levels are substituted for log-transformed testosterone.

Table 3
Results (in standardized regression coefficients) for main analyses.

β=	X → M (path a)	M → Y (path b)	X → Y (path c)	Direct effect (path c')	Indirect effect (path ab)	X, M VIF
			Masculine identity			(Covar VIF)
Model 1 (N 418)	X GenMas	M Testosterone				
Y = PAQ mas	0.68**	0.10	0.19**	0.26** [0.12, 0.38]	0.07 [0.16, 0.02]	1.88
Covar Sex	0.15	0.10	0.23	0.25* [0.07, 0.44]	0.02 [0.05, 0.00]	3.80, 2.33 (4.66)
Model 2 (N 418)	X Testosterone	M GenMas				
Y = PAQ mas	0.68**	0.26**	0.08	0.10 [0.23, 0.04]	0.18** [0.08, 0.27]	1.88
Covar Sex	0.09	0.25*	0.08	0.10 [0.25, 0.03]	0.02 [0.00, 0.05]	2.33, 3.80 (4.66)
Model 3 (N 418)	**R2change**	**β**	**t**			
X*M → PAQ mas	0.01	0.08	1.17			
Covar Sex	0.01	0.10	1.26			
			Feminine identity			
Model 1 (N 421)	X GenFem	M Testosterone				
Y = PAQ fem	0.72**	0.06	0.24**	0.29** [0.15, 0.43]	0.04 [0.14, 0.05]	2.08
Covar Sex	0.20	0.04	0.39**	0.40* [0.16, 0.65]	0.01 [0.04, 0.02]	6.95, 2.31 (7.61)
Model 2 (N 421)	X Testosterone	M GenFem				
Y = PAQ fem	0.72**	0.29**	0.15*	0.06 [0.08, 0.19]	0.21** [0.32, 0.10]	
Covar Sex	0.07	0.40**	0.01	0.04 [0.09, 0.17]	0.03 [0.06, 0.00]	2.31, 6.95 (7.61)
Model 3 (N 421)	**R2change**	**β**	**t**			
X*M → PAQ fem	0.02	0.20*	2.69			
Covar Sex	0.01	0.20	2.47			

Note. Values for models 1 and 2 are standardized betas. Values in brackets are bootstrapped 95 % confidence intervals. *p .01; **p .001. VIF = variance inflation factor.

Table 4
Results (in standardized regression coefficients) for main analyses including only participants assigned female at birth (AFAB).

AFAB	X → M (path a)	M → Y (path b)	X → Y ' (path c)	Direct effect (path c')	Indirect effect (path ab)
Masculine identity					
Model 1 (N 270)	X GenMas	M Testosterone			
Y = PAQ mas	0.14	0.08	0.01	0.02 [0.11, 0.15]	0.01 [0.03, 0.00]
Model 2 (N 270)	X Testosterone	M GenMas			
Y = PAQ mas	0.14	0.02	0.07	0.08 [0.21, 0.04]	0.01 [0.02, 0.02]
Model 3 (N 270)	**R2change**	**β**	**t**		
X*M → PAQ mas	0.02	0.12	2.07		
Feminine identity					
Model 1 (N 273)	X GenFem	M Testosterone			
Y = PAQ fem	0.12	0.01	0.24**	0.24** [0.12, 0.38]	0.01 [0.02, 0.02]
Model 2 (N 273)	X Testosterone	M GenFem			
Y = PAQ fem	0.12	0.24**	0.04	0.01 [0.12, 0.11]	0.03 [0.07, 0.00]
Model 3 (N 273)	**R2change**	**β**	**t**		
X*M → PAQ fem	0.01	0.02	0.31		

Note. Values for models 1 and 2 are standardized betas. Values in brackets are bootstrapped 95 % confidence intervals. *p .01; **p .001.

Table 5
Results (in standardized regression coefficients) for main analyses including only participants assigned male at birth (AMAB).

AMAB	X → M (path a)	M → Y (path b)	X → Y ' (path c)	Direct effect (path c')	Indirect effect (path ab)
Masculine identity					
Model 1 (N 148)	X GenMas	M Testosterone			
Y = PAQ mas	0.01	0.01	0.37**	0.37** [0.22, 0.51]	0.01 [-0.01, 0.01]
Model 2 (N 148)	X Testosterone	M GenMas			
Y = PAQ mas	0.01	0.37**	0.01	0.01 [-0.14, 0.13]	0.01 [-0.06, 0.06]
Model 3 (N 148)	**R2change**	**β**	**t**		
X*M → PAQ mas	0.01	0.01	0.15		
Feminine identity					
Model 1 (N 148)	X GenFem	M Testosterone			
Y = PAQ fem	-0.10	0.14	0.02	0.04 [-0.13, 0.20]	-0.01 [-0.06, 0.01]
Model 2 (N 148)	X Testosterone	M GenFem			
Y = PAQ fem	-0.10	0.04	0.14	0.14 [-0.01, 0.32]	0.01 [-0.03, 0.02]
Model 3 (N 148)	**R2change**	**β**	**t**		
X*M → PAQ fem	0.01	0.07	0.86		

Note. Values for models 1 and 2 are standardized betas. Values in brackets are bootstrapped 95 % confidence intervals. *p .01; **p .001.

women. Overall, contrary to popular belief and the testosterone-masculinity hypothesis, there was no evidence that basal testosterone levels, which are significantly higher in men compared to women, correlated with masculine or feminine stereotyped personality traits. Further, there was no compelling evidence that testosterone explains the relationship between sex/gender identity and these personality traits. Further, when controlling for sex assigned at birth and testing relationships within-sex, there was also generally no supporting evidence that variability in testosterone, above and beyond the binary and within sex, predicts individual differences gender-stereotyped personality traits.

These finding are contrary to early studies of within-sex variation in testosterone and personality (e.g., Baucom et al., 1985), but consistent with more recent large sample and meta-analytic evidence that testosterone (and its interaction with cortisol) is generally unrelated to self-reported personality features, even those thought to be sex/gender-linked (Casto et al., 2023; Grebe et al., 2019a; Sundin et al., 2021). The current study contributes to this literature and offers new evidence that these conclusions hold also for both stereotypically masculine and feminine traits.

The model comparison approach of testing testosterone and a continuous measure of gender identity as alternating mediators provided the opportunity to explore the three-way relationships in depth. Further, repeating these analyses with and without controlling for sex assigned at birth and within the sample of women and men allowed for a test of the testosterone-masculinity hypothesis between, within, and beyond sex assigned at birth. Under these different analyses, some exploratory nuance emerged. When including gender identity as the mediator (Model 2), the small direct association between testosterone and gender-stereotyped personality in the full sample, was statistically explained by the indirect effect of the corresponding gender identity measure. However, when controlling for sex assigned at birth, the relationship between testosterone and gender identity, and indirect effect of gender identity, were substantially reduced and no longer significant. These results collectively suggest that self-perceptions about one's own masculinity and femininity explain the small effects of sex-differences in testosterone on gendered personality traits. Critically, there is not strong evidence that individual variability in basal testosterone is related to these aspects of the self. Importantly, our continuous measure of gender identity remained a moderate predictor of the corresponding gender-stereotyped personality traits above and beyond sex assigned at birth, indicating the utility of self-perceptions of gender along a continuum and beyond the binary for this area of research.

Finally, the moderation models showed no clear evidence that testosterone levels and masculine identity interact to predict masculine attributes. However, there was a small but statistically significant effect that testosterone levels and feminine identity interact to predict feminine attributes. Specifically, in a mixed-sex sample, higher testosterone predicted *higher* scores on *feminine* attributes, but only for those who had a moderate to low rating on feminine gender identity (mostly AMAB individuals). This effect was in the same direction but was weakened when controlling for sex assigned at birth and did not pass a robustness test using raw testosterone levels.

There was another exploratory trend for testosterone that is worth noting: when controlling for sex assigned at birth, there was a small effect, non-significant after correcting for multiple tests, for a positive correlation between testosterone and self-perceived masculinity and negative correlation between testosterone and self-perceived femininity. This may indicate that although basal testosterone was generally unrelated to agency and communality personality attributes, beyond binary sex, variation in basal testosterone may account for some variation in gender identity in terms of perceiving oneself as more or less masculine

or feminine. Further research is required to substantiate this finding.

4.1. Other models of testosterone s effect on sex/gender-typed personality and behavior

One critical characteristic to understanding testosterone's effects on behavior is differentiating between the organizational effects resulting from early exposure in development and the effects of individual- and sex-differences in testosterone levels in adulthood. Results from this study do not speak to whether testosterone exposure in development explains sex/gender effects on personality. Some research on sex differences in spatial cognition, for example, has tested both early exposure and adulthood levels of testosterone as a mediator and found mixed results (Alarcon et al., 2014; Erdmann et al., 2019; Falter et al., 2006; Puts et al., 2010; Toivainen et al., 2018; Vuoksimaa et al., 2012). Although the research designs are challenging and likely require lengthy longitudinal data collection, a test of both fetal and pubertal differences in testosterone exposure on personality attributes would be necessary for a more complete understanding of testosterone's role in promoting sex/gender differences in certain personality traits. Further, such research would also benefit from testing competing models of the role of psychosocial and cultural factors (e.g., Davis and Risman, 2015).

Other models of testosterone's effect on gendered behavior focuses on the importance of within-individual shifts. According to the challenge hypothesis (Wingfield et al., 1990, 2019), an individual's testosterone may increase to a higher set-point during times when mating resources are available (e.g., breeding season) and then increase further in the presence of a potential mate or territorial challenge to promote courtship, dominant, or aggressive behavior that would benefit success in that context (i.e., activational effects of testosterone). Outside of these contexts or following successful competition or mating, testosterone should decrease to reduce risks associated with aggression and promote parental care of offspring. Originally articulated in research with seasonally breeding male birds, support for the challenge hypothesis has also been shown in males and females (or men and women), but often with some notable sex/gender differences (for review, Carre and Archer, 2018; Casto and Edwards, 2016; Gray et al., 2020; Grebe et al., 2022; Grebe et al., 2019b).

Evidence for the importance of within-individual shifts in testosterone in relation to social status-seeking behavior in humans has been shown in studies of endogenous testosterone reactivity to social competition (e.g., Casto et al., 2020), studies of exogenous testosterone administration on subsequent behavior (Geniole et al., 2019; Losecaat Vermeer et al., 2020), and longitudinal correlations between testosterone levels and changes in social standing (e.g., Cheng et al., 2018). Of course, these studies also find that critical contextual and person factors moderate these effects, including sex/gender. This area of research may also benefit from the inclusion of nonbinary indices of gender identification and a consideration of socialized aspects of gender that may impact the expression of status behavior in certain tasks and contexts (Casto and Prasad, 2017).

In addition to within-person shifts in testosterone associated with the social context, testosterone fluctuations due to reproductive physiology may be important for a full understanding of the relationship between testosterone and gendered behavior. In women, there is evidence that testosterone levels in blood are higher at mid-cycle, in association with ovulation, relative to other parts of the menstrual cycle (Atukorala et al., 2022; Bui et al., 2013). Although thought to be linked to much greater increases in estradiol at this time, the mid-cycle ovulatory window is also associated with peaks in self-perceived attractiveness, desirability, and self-esteem (Schleifenbaum et al., 2021). Menstrual cycle shifts in testosterone could be explored in future research in relation to sex differences in personality and self-reports of gender stereotyped attributes. An important contribution of such research could be a greater appreciation of the ways in which sex/gender differences in behavior and personality are not representative of "hard-wired" or fixed effects. Rather, biological processes including hormonal patterns of responding to the environment and those underlying the reproductive system allow for more flexible expressions of the self.

One final consideration for within-individual shifts in testosterone and effects on gendered behavior could be the use of hormonal contraceptives (HCs), which substantially reduce testosterone levels (Casto et al., 2021; Zimmerman et al., 2014). Although HC users do not appear different in gender self-concept in terms of gendered personality attributes or perceptions of masculinity and femininity compared to non-users (Beltz et al., 2019; Nielson and Beltz, 2021), within-individual shifts in testosterone suppression from beginning HC use, or testosterone increases from discontinuing HC use, could be studied in future research. Importantly, factors such as reasons for using HCs, relationship status, and androgenicity of the HC formula should be considered in such research (Beltz, 2022; Beltz et al., 2022).

4.2. Non-binary approaches to sex/gender and personality

A core goal of the present study was to explore considerations of the variable of sex/gender beyond the typical binary categories. Our masculine and feminine gender identity rating scales provided a two-dimensional measure for people to self-identify their masculinity and femininity along a continuum. We found that these continuous measures of masculine and feminine identity predicted their corresponding gendered attribute score above and beyond categorical sex; those with a higher masculine and feminine identity reported higher masculine and feminine stereotyped attributes, respectively. This suggests that gender identity explains some variability in gender-stereotyped personality, independent of sex. Further, within-sex analyses revealed further complexity in this relationship. Specifically, in AFAB individuals, only feminine identity ratings predicted feminine attributes whereas in AMAB individuals, only masculine identity ratings predicted masculine attributes. Overall, these results suggest that a continuous measure for gender identity may be beneficial to future research on sex/gender differences in gender-stereotyped personality traits (for examples with cognitive outcomes, see Kelly and Beltz, 2022; Pletzer et al., 2019).

One potential criticism of determining gender identity through masculinity and femininity ratings is that doing so may require participants to rely on stereotypical thinking and the assumption that individuals have a consistent and divergent concept of the traits of average men and women. For example, some people have highly differentiated and stereotyped views of men and women where the polar ends represent an archetype. Others have a less polarized view of sex/gender and may view the ends of the feminine and masculine scale as the average man or woman. Self-identification with these social groups are also influenced by social categorization processes that place higher or lower, positive or negative value judgements on typical masculinity and femininity (Johnson et al., 2015; Wood and Eagly, 2015). In future research, it may be necessary to assess individual differences in stereotypical thinking about sex/gender and value judgements of these categories prior to or in combination with self-perceptions of masculinity and femininity.

Another drawback of our gender identity measure, which also limits interpretation in the present study, is that it is strongly correlated with sex assigned at birth in our non-gender-diverse sample of college aged participants. Although participants could have indicated their identity as more gender neutral, the results were generally bipolar. Thus, our gender identity variable, as is typical of measures of gender, is conflated with sex assigned at birth. Future research on the topics explored in this study would greatly benefit from the ability to understand non-binary gender identity and better disentangle psychosocial aspects of sex/gender.

Recent theorizing in social psychology has asserted that binary gender perceptions of masculinity and femininity form the basis of a universal two-factor structure of personality (Martin and Slepian, 2021). These authors argue that the pervasiveness of a "Big Two" dimensional

axis of personality, e.g., agency versus communality, combined with the continued ability of gender to predict these factors, suggests that gender is, in fact, the underlying core structure of personality; they are one and the same. Only moderate sex/gender differences in the agency scale (labeled as "masculine") and communal scale (labeled "feminine") were found in the current study and only for some traits; there was substantial overlap between the sexes. Our continuous measure of masculine and feminine gender identification allows for an opportunity to quantify their association with personality factors of agency and communality, which were only low to moderate at best ($r = 0.19$–0.23 for mas-agency and fem-communality, and correlations were reduced when controlling for sex assigned at birth). Thus, at least in our data, gender identity and a two-factor personality structure are certainly not the same thing. These findings contribute to the continued debate about the magnitude and importance of sex differences in personality (Eagly and Revelle, 2022). Again, we advocate that future research in this area consider continuous and non-binary measures of gender identity (e.g., Jacobson and Joel, 2019) in place of, or in addition to, categorical and binary indices of sex/gender.

4.3. Conclusion

In a mixed sex sample of over 400 university undergraduates, there was no compelling evidence that sex-differences in basal testosterone levels predict gender-stereotyped personality attributes. Although masculine-identifying individuals tended to report being slightly more competitive and decisive, for example, and feminine-identifying individuals slightly more emotional and devoted to others, for example, testosterone level did not explain this small to moderate effect of sex/gender on personality. Further, testosterone was unrelated to these attributes when controlling for sex assigned at birth and when testing these relationships within sex assigned at birth. Contributing to a more gender diverse approach to assessing sex/gender relationships with personality and testosterone, our continuous measure of masculine and feminine gender identity predicted additional variance in these outcomes beyond binary sex category and could be an important tool in future research. We discuss limitations of the models we tested for speaking to developmental, situational, and within-individual change effects of testosterone and provide recommendations for future research in this area that considers both the hormonal *and* psychosocial factors of sex/gender. Sex/gender is evident in personality, but these effects may be more likely to arise from a multitude of evolutionary, organizational, and socio-cultural factors rather than individual differences in basal testosterone.

Ethics approval

All study procedures were approved by the university Institutional Review Boards. All persons gave their informed consent prior to their inclusion in the study.

Funding

This research was supported by National Science Foundation Grant# 1451848 awarded to Pranjal Mehta; a National Research Council Fellowship awarded to Kathleen Casto; and a Columbia University Provost grant awarded to Modupe Akinola.

CRediT authorship contribution statement

Kathleen V. Casto: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Dale J. Cohen:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Modupe Akinola:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Pranjal H. Mehta:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors have no relevant financial or non-financial interests to disclose.

Data availability

The data and R code for this study are available at the Open Science Framework repository: https://osf.io/bny4j/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yhbeh.2024.105540.

References

Alarcon, G., Cservenka, A., Fair, D.A., Nagel, B.J., 2014. Sex differences in the neural substrates of spatial working memory during adolescence are not mediated by endogenous testosterone. Brain Res. 1593, 40–54. https://doi.org/10.1016/j.brainres.2014.09.057.

Alvergne, A., Jokela, M., Faurie, C., Lummaa, V., 2010. Personality and testosterone in men from a high-fertility population. Personal. Individ. Differ. 49 (8), 840–844. https://doi.org/10.1016/j.paid.2010.07.006.

Archer, J., 2019. The reality and evolutionary significance of human psychological sex differences. Biol. Rev. 94 (4), 1381–1415. https://doi.org/10.1111/brv.12507.

Armstrong, T.A., Boisvert, D.L., Wells, J., Lewis, R.H., Cooke, E.M., Woeckener, M., Kavish, N., Vietto, N., Harper, J.M., 2022. Testosterone, cortisol, and criminal behavior in men and women. Horm. Behav. 146, 105260 https://doi.org/10.1016/j.yhbeh.2022.105260.

Arnold, A.P., 2009. The organizational–activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. Horm. Behav. 55 (5), 570–578. https://doi.org/10.1016/j.yhbeh.2009.03.011.

Atukorala, K.R., Silva, W., Amarasiri, L., Fernando, D., 2022. Changes in serum testosterone during the menstrual cycle – an integrative systematic review of published literature. Gynecol. Reprod. Endocrinol. Metab. 3 (1), 9–20.

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., Hackett, G., Hines, M., 2009. Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. Psychol. Sci. 20 (2), 144–148. https://doi.org/10.1111/j.1467-9280.2009.02279.x.

Bahr, R., 1992. The Virility Factor: Masculinity Through Testosterone, the Male Sex Hormone. Factor Press.

Bandini, E., Corona, G., Ricca, V., Fisher, A.D., Lotti, F., Sforza, A., Faravelli, C., Forti, G., Mannucci, E., Maggi, M., 2009. Hysterical traits are not from the uterus but from the testis: a study in men with sexual dysfunction. J. Sexual Med. 6 (8), 2321–2331. https://doi.org/10.1111/j.1743-6109.2009.01322.x.

Banks, T., Dabbs, J.M., 1996. Salivary testosterone and cortisol in a delinquent and violent urban subculture. J. Soc. Psychol. 136 (1), 49–56. https://doi.org/10.1080/00224545.1996.9923028.

Baucom, D.H., Besch, P.K., Callahan, S., 1985. Relation between testosterone concentration, sex role identity, and personality among females. J. Pers. Soc. Psychol. 48 (5), 1218–1226. https://doi.org/10.1037/0022-3514.48.5.1218.

Beltz, A.M., 2022. Hormonal contraceptive influences on cognition and psychopathology: past methods, present inferences, and future directions. Front. Neuroendocrinol. 67, 101037 https://doi.org/10.1016/j.yfrne.2022.101037.

Beltz, A.M., Loviska, A.M., Kelly, D., 2019. No personality differences between oral contraceptive users and naturally cycling women: implications for research on sex hormones. Psychoneuroendocrinology 100, 127–130. https://doi.org/10.1016/j.psyneuen.2018.09.034.

Beltz, A.M., Loviska, A.M., Kelly, D.P., Nielson, M.G., 2022. The link between masculinity and spatial skills is moderated by the estrogenic and progestational activity of oral contraceptives. Front. Behav. Neurosci. 15. https://www.frontiersin.org/articles/10.3389/fnbeh.2021.777911.

Berenbaum, S.A., 1999. Effects of early androgens on sex-typed activities and interests in adolescents with congenital adrenal hyperplasia. Horm. Behav. 35 (1), 102–110. https://doi.org/10.1006/hbeh.1998.1503.

Berenbaum, S.A., Beltz, A.M., 2011. Sexual differentiation of human behavior: effects of prenatal and pubertal organizational hormones. Front. Neuroendocrinol. 32 (2), 183–200. https://doi.org/10.1016/j.yfrne.2011.03.001.

Berenbaum, S.A., Beltz, A.M., 2021. Evidence and implications from a natural experiment of prenatal androgen effects on gendered behavior. Curr. Dir. Psychol. Sci. 30 (3), 202–210. https://doi.org/10.1177/0963721421998341.

Berenbaum, S.A., Hines, M., 1992. Early androgens are related to childhood sex-typed toy preferences. Psychol. Sci. 3 (3), 203–206. https://doi.org/10.1111/j.1467-9280.1992.tb00028.x.

Bhatia, N., Bhatia, S., 2021. Changes in gender stereotypes over time: a computational analysis. Psychol. Women Q. 45 (1), 106–125. https://doi.org/10.1177/0361684320977178.

Briley, D.A., Tucker-Drob, E.M., 2014. Genetic and environmental continuity in personality development: a meta-analysis. Psychol. Bull. 140 (5), 1303–1331. https://doi.org/10.1037/a0037091.

Bui, H.N., Sluss, P.M., Blincko, S., Knol, D.L., Blankenstein, M.A., Heijboer, A.C., 2013. Dynamics of serum testosterone during the menstrual cycle evaluated by daily measurements with an ID-LC–MS/MS method and a 2nd generation automated immunoassay. Steroids 78 (1), 96–101. https://doi.org/10.1016/j.steroids.2012.10.010.

Bullock, J.G., Green, D.P., 2021. The failings of conventional mediation analysis and a design-based alternative. Adv. Methods Pract. Psychol. Sci. 4 (4) https://doi.org/10.1177/25152459211047227, 25152459211047227.

Bullock, J.G., Green, D.P., Ha, S.E., 2010. Yes, but what's the mechanism? (Don't expect an easy answer). J. Pers. Soc. Psychol. 98 (4), 550–558. https://doi.org/10.1037/a0018933.

Carre, J.M., Archer, J., 2018. Testosterone and human behavior: the role of individual and contextual variables. Curr. Opin. Psychol. 19, 149–153. https://doi.org/10.1016/j.copsyc.2017.03.021.

Casto, K.V., Carre, J.M., 2023. Testosterone, sex, and sport. Policy Insights Behav. Brain Sci. 10 (2), 193–200. https://doi.org/10.1177/23727322231197583.

Casto, K.V., Edwards, D.A., 2016. Testosterone, cortisol, and human competition. Horm. Behav. 82, 21–37. https://doi.org/10.1016/j.yhbeh.2016.04.004.

Casto, K.V., Prasad, S., 2017. Recommendations for the study of women in hormones and competition research. Horm. Behav. 92, 190–194. https://doi.org/10.1016/j.yhbeh.2017.05.009.

Casto, K.V., Edwards, D.A., Akinola, M., Davis, C., Mehta, P.H., 2020. Testosterone reactivity to competition and competitive endurance in men and women. Horm. Behav. 123, 104665 https://doi.org/10.1016/j.yhbeh.2019.104665.

Casto, K.V., Arthur, L.C., Hamilton, D.K., Edwards, D.A., 2021. Testosterone, athletic context, oral contraceptive use, and competitive persistence in women. Adapt. Hum. Behav. Physiol. https://doi.org/10.1007/s40750-021-00180-6.

Casto, K.V., Prasad, S., Josephs, R.A., Zilioli, S., Welker, K., Maslov, A., Jones, A.C., Mehta, P.H., 2023. No compelling evidence that self-reported personality traits explain basal testosterone and cortisol's associations with status-relevant behavior. Adapt. Hum. Behav. Physiol. 9 (1), 88–122. https://doi.org/10.1007/s40750-023-00210-5.

Chafkin, J.E., O'Brien, J.M., Medrano, F.N., Lee, H.Y., Yeager, D.S., Josephs, R.A., 2022. Chemiluminescent immunoassay overestimates hormone concentrations and obscures testosterone sex differences relative to LC-MS/MS in a field study of diverse adolescents. Comprehensive Psychoneuroendocrinol. 10, 100132 https://doi.org/10.1016/j.cpnec.2022.100132.

Cheng, J.T., Kornienko, O., Granger, D.A., 2018. Prestige in a large-scale social group predicts longitudinal changes in testosterone. J. Pers. Soc. Psychol. 114 (6), 924–944. https://doi.org/10.1037/pspi0000126.

Choi, N., Fuqua, D.R., Newman, J.L., 2007. Hierarchical confirmatory factor analysis of the Bem sex role inventory. Educ. Psychol. Meas. 67 (5), 818–832. https://doi.org/10.1177/0013164406299106.

Clausen, S., Tajadura-Jimenez, A., Janssen, C. P., & Bianchi-Berthouze, N. (2021). Action sounds informing own body perception influence gender identity and social cognition. Front. Hum. Neurosci., 15. https://www.frontiersin.org/articles/10.3389/fnhum.2021.688170.

Collaer, M.L., Hines, M., 1995. Human behavioral sex differences: a role for gonadal hormones during early development? Psychol. Bull. 118 (1), 55–107. https://doi.org/10.1037/0033-2909.118.1.55.

Constantinescu, M., Moore, D.S., Johnson, S.P., Hines, M., 2018. Early contributions to infants' mental rotation abilities. Dev. Sci. 21 (4), e12613 https://doi.org/10.1111/desc.12613.

Conway, M., Vartanian, L.R., 2000. A status account of gender stereotypes: beyond communality and agency. Sex Roles 43 (3), 181–199. https://doi.org/10.1023/A:1007076813819.

Crewther, B.T., Hecht, M., Potts, N., Kilduff, L.P., Drawer, S., Cook, C.J., 2020. A longitudinal investigation of bidirectional and time-dependent interrelationships between testosterone and training motivation in an elite rugby environment. Horm. Behav. 126, 104866 https://doi.org/10.1016/j.yhbeh.2020.104866.

Cumming, G., 2014. The new statistics: why and how. Psychol. Sci. 25 (1), 7–29. https://doi.org/10.1177/0956797613504966.

Cunningham, E., Benítez, M.E., 2024. From pathology to pleasure: reframing mechanistic studies on same-sex sexual behavior in primates. Horm. Behav. 160, 105476 https://doi.org/10.1016/j.yhbeh.2024.105476.

Dabbs, J.M., Morris, R., 1990. Testosterone, social class, and antisocial behavior in a sample of 4,462 men. Psychol. Sci. 1 (3), 209–211. https://doi.org/10.1111/j.1467-9280.1990.tb00200.x.

Davis, S.N., Risman, B.J., 2015. Feminists wrestle with testosterone: hormones, socialization and cultural interactionism as predictors of women's gendered selves. Soc. Sci. Res. 49, 110–125. https://doi.org/10.1016/j.ssresearch.2014.07.012.

Del Giudice, M., 2022. Measuring sex differences and similarities. In: VanderLaan, D.P., Wong, W.I. (Eds.), Gender and Sexuality Development: Contemporary Theory and Research. Springer International Publishing, pp. 1–38. https://doi.org/10.1007/978-3-030-84273-4_1.

Del Giudice, M., 2023. Ideological bias in the psychology of sex and gender. In: Frisby, C. L., Redding, R.E., O'Donohue, W.T., Lilienfeld, S.O. (Eds.), Ideological and Political Bias in Psychology: Nature, Scope, and Solutions. Springer International Publishing, pp. 743–778. https://doi.org/10.1007/978-3-031-29148-7_28.

Donnelly, K., Twenge, J.M., 2017. Masculine and feminine traits on the Bem sex-role inventory, 1993–2012: a cross-temporal meta-analysis. Sex Roles 76 (9), 556–565. https://doi.org/10.1007/s11199-016-0625-y.

Eagly, A.H., Revelle, W., 2022. Understanding the magnitude of psychological differences between women and men requires seeing the forest and the trees. Perspect. Psychol. Sci. 17 (5), 1339–1358. https://doi.org/10.1177/17456916211046006.

Eagly, A.H., Wood, W., 2013. The nature–nurture debates: 25 years of challenges in understanding the psychology of gender. Perspect. Psychol. Sci. 8 (3), 340–357. https://doi.org/10.1177/1745691613484767.

Egan, S.K., Perry, D.G., 2001. Gender identity: a multidimensional analysis with implications for psychosocial adjustment. Dev. Psychol. 37 (4), 451–463. https://doi.org/10.1037/0012-1649.37.4.451.

Ehrenkranz, J., Bliss, E., Sheard, M.H., 1974. Plasma testosterone: correlation with aggressive behavior and social dominance in man. Psychosom. Med. 36 (6), 469.

Erdmann, K., Schaal, N.K., Meinlschmidt, G., Tegethoff, M., Fröhlich, S., Kozlowski, P., Rivet, N., Jamey, C., Reix, N., Kintz, P., Raul, J.-S., Heil, M., 2019. Sex specific relationships between infants' mental rotation ability and amiotic sex hormones. Neurosci. Lett. 707, 134298 https://doi.org/10.1016/j.neulet.2019.134298.

Falter, C.M., Arroyo, M., Davis, G.J., 2006. Testosterone: activation or organization of spatial cognition? Biol. Psychol. 73 (2), 132–140. https://doi.org/10.1016/j.biopsycho.2006.01.011.

Gaa, J.P., Liberman, D., Edwards, T.A., 1979. A comparative factor analysis of the bem sex role inventory and the personality attributes questionnaire. J. Clin. Psychol. 35 (3), 592–598. https://doi.org/10.1002/1097-4679(197907)35:3 592::AID-JCLP2270350322>3.0.CO;2-W.

Garcia-Sifuentes, Y., Maney, D.L., 2021. Reporting and misreporting of sex differences in the biological sciences. eLife 10, e70817. https://doi.org/10.7554/eLife.70817.

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., Carre, J.M., 2019. Using a psychopharmacogenetic approach to identify the pathways through which—and the people for whom—testosterone promotes aggression. Psychol. Sci. 30 (4), 481–494. https://doi.org/10.1177/0956797619826970.

Giudice, M., Booth, T., Irwing, P., 2012. The distance between Mars and Venus: measuring global sex differences in personality. PloS One 7 (1), e29265. https://doi.org/10.1371/journal.pone.0029265.

Gray, P.B., Straftis, A.A., Bird, B.M., McHale, T.S., Zilioli, S., 2020. Human reproductive behavior, life history, and the challenge hypothesis: a 30-year review, retrospective and future directions. Horm. Behav. 123, 104530 https://doi.org/10.1016/j.yhbeh.2019.04.017.

Grebe, N.M., Del Giudice, M., Emery Thompson, M., Nickels, N., Ponzi, D., Zilioli, S., Maestripieri, D., Gangestad, S.W., 2019a. Testosterone, cortisol, and status-striving personality features: a review and empirical evaluation of the dual hormone hypothesis. Horm. Behav. 109, 25–37. https://doi.org/10.1016/j.yhbeh.2019.01.006.

Grebe, N.M., Sarafin, R.E., Strenth, C.R., Zilioli, S., 2019b. Pair-bonding, fatherhood, and the role of testosterone: a meta-analytic review. Neurosci. Biobehav. Rev. 98, 221–233. https://doi.org/10.1016/j.neubiorev.2019.01.010.

Grebe, N.M., Sheikh, A., Drea, C.M., 2022. Integrating the female masculinization and challenge hypotheses: female dominance, male deference, and seasonal hormone fluctuations in adult blue-eyed black lemurs (Eulemur flavifrons). Horm. Behav. 139, 105108 https://doi.org/10.1016/j.yhbeh.2022.105108.

Hausmann, M., Schoofs, D., Rosenthal, H.E.S., Jordan, K., 2009. Interactive effects of sex hormones and gender stereotypes on cognitive sex differences—a psychobiosocial approach. Psychoneuroendocrinology 34 (3), 389–401. https://doi.org/10.1016/j.psyneuen.2008.09.019.

Hayes, A.F., 2020. Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach. Guilford Publications.

Hayes, A.F., Scharkow, M., 2013. The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: does method really matter? Psychol. Sci. 24 (10), 1918–1927. https://doi.org/10.1177/0956797613480187.

Helmreich, R.L., Spence, J.T., Wilhelm, J.A., 1981. A psychometric analysis of the personal attributes questionnaire. Sex Roles 7 (11), 1097–1108. https://doi.org/10.1007/BF00287587.

Hentschel, T., Heilman, M.E., Peus, C.V., 2019. The multiple dimensions of gender stereotypes: a current look at men's and women's characterizations of others and themselves. Front. Psychol. 10, 11. https://doi.org/10.3389/fpsyg.2019.00011.

Hill, S.A., Fekken, G.C., Bond, S.L., 2000. Factor structure integrity of the Personal Attributes Questionnaire: an English–French comparison. Can. J. Behav. Sci./Revue Canadienne Des Sciences Du Comportement 32 (4), 234–242. https://doi.org/10.1037/h0087120.

Hines, M., 2002. 70—sexual differentiation of human brain and behavior. In: Pfaff, D.W., Arnold, A.P., Fahrbach, S.E., Etgen, A.M., Rubin, R.T. (Eds.), Hormones, Brain and Behavior. Academic Press, pp. 425–462. https://doi.org/10.1016/B978-012532104-4/50072-X.

Hines, M., Golombok, S., Rust, J., Johnston, K.J., Golding, J., Parents and Children Study Team, Avon Longitudinal Study of, 2002. Testosterone during pregnancy and gender role behavior of preschool children: a longitudinal, population study. Child Develop. 73 (6), 1678–1687. https://doi.org/10.1111/1467-8624.00498.

Hopwood, C.J., Donnellan, M.B., Blonigen, D.M., Krueger, R.F., McGue, M., Iacono, W. G., Burt, S.A., 2011. Genetic and environmental influences on personality trait stability and growth during the transition to adulthood: a three wave longitudinal study. J. Pers. Soc. Psychol. 100 (3), 545–556. https://doi.org/10.1037/a0022409.

Hyde, J.S., Bigler, R.S., Joel, D., Tate, C.C., van Anders, S.M., 2019. The future of sex and gender in psychology: five challenges to the gender binary. Am. Psychol. 74 (2), 171–193. https://doi.org/10.1037/amp0000307.

Jacobson, R., Joel, D., 2019. Self-reported gender identity and sexuality in an online sample of cisgender, transgender, and gender-diverse individuals: an exploratory study. J. Sex Res. 56 (2), 249–263. https://doi.org/10.1080/00224499.2018.1523998.

Johnson, K.L., Lick, D.J., Carpinella, C.M., 2015. Emergent research in social vision: an integrated approach to the determinants and consequences of social categorization. Soc. Personal. Psychol. Compass 9 (1), 15–30. https://doi.org/10.1111/spc3.12147.

Jordan-Young, R., Karkazis, K., 2019. Testosterone: An Unauthorized Biography. Havard University Press.

Kaiser, T., Del Giudice, M., Booth, T., 2020. Global sex differences in personality: replication with an open online dataset. J. Pers. 88 (3), 415–429. https://doi.org/10.1111/jopy.12500.

Kelly, D.P., Beltz, A.M., 2022. Spatial skills and Self-perceived masculinity: considering college major STEM-ness, directionality, and gender. Sex Roles 87 (5), 251–266. https://doi.org/10.1007/s11199-022-01308-y.

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. Horm. Behav. 123, 104657 https://doi.org/10.1016/j.yhbeh.2019.104657.

Losecaat Vermeer, A.B., Krol, I., Gausterer, C., Wagner, B., Eisenegger, C., Lamm, C., 2020. Exogenous testosterone increases status-seeking motivation in men with unstable low social status. Psychoneuroendocrinology 113, 104552. https://doi.org/10.1016/j.psyneuen.2019.104552.

Maney, D., 2016. Perils and pitfalls of reporting sex differences. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371 (20150119) https://doi.org/10.1098/rstb.2015.0119.

Maney, D.L., Rich-Edwards, J.W., 2023. Sex-inclusive biomedicine: are new policies increasing rigor and reproducibility? Womens Health Issues 33 (5), 461–464. https://doi.org/10.1016/j.whi.2023.03.004.

Martin, A.E., Slepian, M.L., 2021. The primacy of gender: gendered cognition underlies the big two dimensions of social cognition. Perspect. Psychol. Sci. 16 (6), 1143–1158. https://doi.org/10.1177/1745691620904961.

Mazur, A., 2016. Testosterone is high among young Black men with little education. Front. Sociol. 1. https://www.frontiersin.org/articles/10.3389/fsoc.2016.00001.

McCarthy, M.M., Wright, C.L., Schwarz, J.M., 2009. New tricks by an old dogma: mechanisms of the organizational/activational hypothesis of steroid-mediated sexual differentiation of brain and behavior. Horm. Behav. 55 (5), 655–665. https://doi.org/10.1016/j.yhbeh.2009.02.012.

McCreary, D.R., Steinberg, M., 1992. The Personal Attributes Questionnaire in Britain: establishing construct validity. Br. J. Soc. Psychol. 31 (4), 369–378. https://doi.org/10.1111/j.2044-8309.1992.tb00979.x.

Mehta, P.H., Josephs, R.A., 2010. Testosterone and cortisol jointly regulate dominance: evidence for a dual-hormone hypothesis. Horm. Behav. 58 (5), 898–906. https://doi.org/10.1016/j.yhbeh.2010.08.020.

Morris, J.A., Jordan, C.L., Breedlove, S.M., 2004. Sexual differentiation of the vertebrate nervous system. Nat. Neurosci. 7 (10), 10 https://doi.org/10.1038/nn1325.

Muehlenhard, C.L., Peterson, Z.D., 2011. Distinguishing between sex and gender: History, current conceptualizations, and implications. Sex roles 64, 791–803. https://doi.org/10.1007/s11199-011-9932-5.

Nielson, M.G., Beltz, A.M., 2021. Oral contraceptive use is not related to gender self-concept. Psychoneuroendocrinology 129 (105271). https://doi.org/10.1016/j.psyneuen.2021.105271.

Nordenström, A., Servin, A., Bohlin, G., Larsson, A., Wedell, A., 2002. Sex-typed toy play behavior correlates with the degree of prenatal androgen exposure assessed by CYP21 genotype in girls with congenital adrenal hyperplasia. J. Clin. Endocrinol. Metabol. 87 (11), 5119–5124. https://doi.org/10.1210/jc.2001-011531.

Patterson, M.M., 2012. Self-perceived gender typicality, gender-typed attributes, and gender stereotype endorsement in elementary-school-aged children. Sex Roles 67 (7), 422–434. https://doi.org/10.1007/s11199-012-0184-9.

Phoenix, C.H., Goy, R.W., Gerll, A.A., Young, W.C., 1959. Organizing action of prenatally administered testostreone propionate on the tissues mediating mating behavior in the female guinea pig. Endocrinology 65 (3), 369–382. https://doi.org/10.1210/endo-65-3-369.

Pletzer, B., Petasis, O., Ortner, T., Cahill, L., 2015. Interactive effects of culture and sex hormones on the sex role self-concept. Front. Neurosci. 9. https://www.frontiersin.org/articles/10.3389/fnins.2015.00240.

Pletzer, B., Steinbeisser, J., van Laak, L., Harris, T., 2019. Beyond biological sex: interactive effects of gender role and sex hormones on spatial abilities. Front. Neurosci. 13. https://www.frontiersin.org/articles/10.3389/fnins.2019.00675.

Puts, D., 2016. Human sexual selection. Curr. Opin. Psychol. 7, 28–32. https://doi.org/10.1016/j.copsyc.2015.07.011.

Puts, D.A., Cardenas, R.A., Bailey, D.H., Burriss, R.P., Jordan, C.L., Breedlove, S.M., 2010. Salivary testosterone does not predict mental rotation performance in men or women. Horm. Behav. 58 (2), 282–289. https://doi.org/10.1016/j.yhbeh.2010.03.005.

Rodrigues dos Santos, M., Bhasin, S., 2021. Benefits and risks of testosterone treatment in men with age-related decline in testosterone. Annu. Rev. Med. 72 (1), 75–91. https://doi.org/10.1146/annurev-med-050219-034711.

Rohrer, J.M., Hünermund, P., Arslan, R.C., Elson, M., 2022. That's a lot to process! Pitfalls of popular path models. Adv. Methods Pract. Psychol. Sci. 5 (2) https://doi.org/10.1177/25152459221095827, 25152459221095827.

Schleifenbaum, L., Driebe, J.C., Gerlach, T.M., Penke, L., Arslan, R.C., 2021. Women feel more attractive before ovulation: evidence from a large-scale online diary study. Evol. Human Sci. 3, e47 https://doi.org/10.1017/ehs.2021.44.

Schmitt, D.P., Long, A.E., McPhearson, A., O'Brien, K., Remmert, B., Shah, S.H., 2017. Personality and gender differences in global perspective. Int. J. Psychol. 52 (S1), 45–56. https://doi.org/10.1002/ijop.12265.

Sellers, J.G., Mehl, M.R., Josephs, R.A., 2007. Hormones and personality: testosterone as a marker of individual differences. J. Res. Pers. 41 (1), 126–138. https://doi.org/10.1016/j.jrp.2006.02.004.

Servin, A., Nordenström, A., Larsson, A., Bohlin, G., 2003. Prenatal androgens and gender-typed behavior: a study of girls with mild and severe forms of congenital adrenal hyperplasia. Dev. Psychol. 39 (3), 440–450. https://doi.org/10.1037/0012-1649.39.3.440.

Shirazi, T.N., Self, H., Rosenfield, K.A., Dawood, K., Welling, L.L.M., Cardenas, R., Bailey, J.M., Balasubramanian, R., Delaney, A., Breedlove, S.M., Puts, D.A., 2022. Low perinatal androgens predict recalled childhood gender nonconformity in men. Psychol. Sci. 33 (3), 343–353. https://doi.org/10.1177/09567976211036075.

Smeets-Janssen, M.M.J., Roelofs, K., Pelt, J. van, Spinhoven, P., Zitman, F.G., Penninx, B. W.J.H., Giltay, E.J., 2015. Salivary testosterone is consistently and positively associated with extraversion: results from the Netherlands study of depression and anxiety. Neuropsychobiology 71 (2), 76–84. https://doi.org/10.1159/000369024.

Snyder, P.J., Bhasin, S., Cunningham, G.R., Matsumoto, A.M., Stephens-Shields, A.J., Cauley, J.A., Gill, T.M., Barrett-Connor, E., Swerdloff, R.S., Wang, C., Ensrud, K.E., Lewis, C.E., Farrar, J.T., Cella, D., Rosen, R.C., Pahor, M., Crandall, J.P., Molitch, M. E., Resnick, S.M., Ellenberg, S.S., 2018. Lessons from the testosterone trials. Endocr. Rev. 39 (3), 369–386. https://doi.org/10.1210/er.2017-00234.

Spence, J.T., Helmreich, R.L., 1978. Masculinity and Femininity: Their Psychological Dimensions, Correlates, and Antecedents. University of Texas Press. https://doi.org/10.7560/764439.

Sundin, Z.W., Chopik, W.J., Welker, K.M., Ascigil, E., Brandes, C.M., Chin, K., Ketay, S., Knight, E.L., Kordsmeyer, T.L., McLarney-Vesotski, A.R., Prasad, S., Reese, Z.A., Roy, A.R.K., Sim, L., Stern, J., Carre, J.M., Edelstein, R.S., Mehta, P.H., Penke, L., Tackett, J.L., 2021. Estimating the associations between big five personality traits, testosterone, and cortisol. Adapt. Hum. Behav. Physiol. 7 (3), 307–340. https://doi.org/10.1007/s40750-020-00159-9.

Swift-Gallant, A., Aung, T., Rosenfield, K., Dawood, K., Puts, D., 2023. Organizational effects of gonadal hormones on human sexual orientation. Adapt. Hum. Behav. Physiol. 9 (4), 344–370. https://doi.org/10.1007/s40750-023-00226-x.

Tacikowski, P., Fust, J., Ehrsson, H.H., 2020. Fluidity of gender identity induced by illusory body-sex change. Sci. Rep. 10 (1), 1 https://doi.org/10.1038/s41598-020-71467-z.

Tibbe, T.D., Montoya, A.K., 2022. Correcting the bias correction for the bootstrap confidence interval in mediation analysis. Front. Psychol. 13, 810258 https://doi.org/10.3389/fpsyg.2022.810258.

Toivainen, T., Pannini, G., Papageorgiou, K.A., Malanchini, M., Rimfeld, K., Shakeshaft, N., Kovas, Y., 2018. Prenatal testosterone does not explain sex differences in spatial ability. Sci. Rep. 8 (1), 1 https://doi.org/10.1038/s41598-018-31704-y.

van Anders, S.M., 2013. Beyond masculinity: testosterone, gender/sex, and human social behavior in a comparative context. Front. Neuroendocrinol. 34 (3), 198–210. https://doi.org/10.1016/j.yfrne.2013.07.001.

Vukasovic, T., Bratko, D., 2015. Heritability of personality: a meta-analysis of behavior genetic studies. Psychol. Bull. 141 (4), 769–785. https://doi.org/10.1037/bul0000017.

Vuoksimaa, E., Kaprio, J., Eriksson, C.J.P., Rose, R.J., 2012. Pubertal testosterone predicts mental rotation performance of young adult males. Psychoneuroendocrinology 37 (11), 1791–1800. https://doi.org/10.1016/j.psyneuen.2012.03.013.

Wallen, K., 2005. Hormonal influences on sexually differentiated behavior in nonhuman primates. Front. Neuroendocrinol. 26 (1), 7–26. https://doi.org/10.1016/j.yfrne.2005.02.001.

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. https://doi.org/10.1016/j.psyneuen.2016.05.022.

Williams, J.S., Fattori, M.R., Honeyborne, I.R., Ritz, S.A., 2023. Considering hormones as sex- and gender-related factors in biomedical research: challenging false dichotomies and embracing complexity. Horm. Behav. 156, 105442 https://doi.org/10.1016/j.yhbeh.2023.105442.

Wingfield, J.C., Hegner, R.E., Dufty, Alfred M., Ball, G.F., 1990. The "challenge hypothesis": theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. Am. Nat. 136 (6), 829–846. https://doi.org/10.1086/285134.

Wingfield, J.C., Goymann, W., Jalabert, C., Soma, K.K., 2019. Concepts derived from the challenge hypothesis. Horm. Behav. 115, 104550 https://doi.org/10.1016/j.yhbeh.2019.06.014.

Woitowich, N.C., Beery, A., Woodruff, T., 2020. A 10-year follow-up study of sex inclusion in the biological sciences. eLife 9, e56344. https://doi.org/10.7554/eLife.56344.

Wood, W., Eagly, A., 2012. Biosocial construction of sex differences and similarities in behavior. In: Advances in Experimental Social Psychology, 46, pp. 55–123. https://doi.org/10.1016/B978-0-12-394281-4.00002-7.

Wood, W., Eagly, A.H., 2015. Two traditions of research on gender identity. Sex Roles 73 (11), 461–473. https://doi.org/10.1007/s11199-015-0480-2.

Zimmerman, Y., Eijkemans, M.J.C., Coelingh Bennink, H.J.T., Blankenstein, M.A., Fauser, B.C.J.M., 2014. The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis. Hum. Reprod. Update 20 (1), 76–105. https://doi.org/10.1093/humupd/dmt038.