

# A Case of Evolutionary Mismatch? Why Facial Width-to-Height Ratio May Not Predict Behavioral Tendencies

Psychological Science  
2019, Vol. 30(7) 1074–1081  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0956797619849928  
www.psychologicalscience.org/PS  


**Dawei Wang<sup>1</sup>, Krishnan Nair<sup>1,2</sup>, Maryam Kouchaki<sup>1</sup>,  
Edward J. Zajac<sup>1</sup>, and Xiuxi Zhao<sup>3</sup>**

<sup>1</sup>Kellogg School of Management, Northwestern University; <sup>2</sup>Rotterdam School of Management, Erasmus University; and <sup>3</sup>Center for Creative Leadership, Singapore

## Abstract

This study contributes to the growing literature linking physical characteristics and behavioral tendencies by advancing the current debate on whether a person's facial width-to-height ratio (fWHR) predicts a variety of antisocial tendencies. Specifically, our large-scale study avoided the social-desirability bias found in self-reports of behavioral tendencies by capturing survey data not only from more than 1,000 business executives but also from evaluators who reported knowing the focal individuals well. With this improved research design, and after conducting a variety of analyses, we found very little evidence of fWHR predicting antisocial tendencies. In light of prior research linking fWHR to social perceptions of evaluators, our results are suggestive of an evolutionary mismatch, whereby a physical characteristic once tied to antisocial tendencies in ancestral environments is—in modern environments—not predictive of such behaviors but instead predictive of biased perceptions.

## Keywords

evolutionary psychology, facial features, physical appearance, social behavior

Received 6/12/18; Revision accepted 4/21/19

As part of the burgeoning body of evolutionary-psychology research linking physical characteristics and behavioral tendencies, a consensus appears to have emerged suggesting that facial width-to-height ratio (fWHR) predicts a host of antisocial tendencies, including threat behavior (Geniole, Denson, Dixson, Carré, & McCormick, 2015), deception and exploitation (Geniole, Keyes, Carré, & McCormick, 2014; Haselhuhn & Wong, 2012; Stirrat & Perrett, 2010), trait dominance (Carré & McCormick, 2008), physical aggression (Goetz et al., 2013), and overall psychopathy (Anderl et al., 2016), and that this is especially so for males. An evolutionary explanation emphasizes that in violent ancestral environments, individuals with a high fWHR were more protected from fatal blows to the face and thus more likely to prevail in physical altercations (Stirrat, Stulp, & Pollet, 2012). This suggests further that males with a high fWHR were more effective ancestrally in garnering influence and access to resources through

the use of threat and intimidation, which led to the development of psychological mechanisms that calibrate antisocial tendencies to fWHR. Evidence showing the greater relevance of fWHR in predicting male behavior can be explained in terms of males disproportionately representing both the perpetrators and victims of violent homicides in cultures across time and space (Daly & Wilson, 1988). Researchers have also proposed that the relationship between fWHR and antisocial tendencies is mediated by testosterone (Carré & McCormick, 2008), but this prediction has received only mixed support (Bird et al., 2016; Lefevre, Lewis, Perrett, & Penke, 2013).

---

## Corresponding Author:

Dawei Wang, Kellogg School of Management, Northwestern University, 2211 Campus Drive, 5100 Wing, Evanston, IL 60208  
E-mail: david.wang@kellogg.northwestern.edu

Although various studies have reported a positive relationship between fWHR and antisocial tendencies, a recent large-scale study by Kosinski (2017) has called into question this growing consensus. Kosinski's study is noteworthy because it found very little evidence linking fWHR and self-reported behavioral tendencies (e.g., cooperativeness, militarism, trustworthiness, sympathy, and morality), and what little evidence was found was also generally stronger for females than for males. A particularly compelling feature of Kosinski's work is that its nonfindings were obtained using a large real-world sample, whereas prior work (mostly laboratory-based) has typically relied on a small number of participants. Indeed, Kosinski speculated that the frequency of marginally significant  $p$  values observed in prior published studies on fWHR and behavioral tendencies may reflect degrees of freedom and file-drawer problems (Nelson, Simmons, & Simonsohn, 2018).

Kosinski's (2017) contribution is undoubtedly important in both highlighting an alternative position on the relevance of fWHR for behavioral tendencies and also identifying specific limitations of prior empirical research in this area. If fWHR does not predict behavioral tendencies, but only differences in social judgments of evaluators who are unfamiliar with targets (as found in prior research; Deska, Lloyd, & Hugenberg, 2018a, 2018b; Efferson & Vogt, 2013; Geniole, MacDonell, & McCormick, 2017; Lefevre & Lewis, 2014), this would open the possibility of an evolutionary mismatch (Li, van Vugt, & Colarelli, 2018). In other words, because the modern world differs in important ways from the violent ancestral environments in which human psychological mechanisms developed, social judgments formed on the basis of fWHR, which may have been adaptive in the evolutionary past, may no longer be accurate.

However, we suggest that before advancing such a concept with confidence, it is necessary to first address an additional important but underexamined limitation found in some of the studies focusing on the link between fWHR and behavioral tendencies (including Kosinski, 2017): namely, a reliance on self-reports and their attendant *social-desirability bias*, the tendency to provide socially acceptable answers (see the discussion of Kosinski's results by Eisenbruch, Lukaszewski, Simmons, Arai, & Roney, 2018). Indeed, some researchers suggest that this bias can explain the persistently low or even negative correlation between self-reported versus other-reported ratings of behavioral tendencies (Atkins & Wood, 2002).

Therefore, we sought in this study to extend the current understanding of the relationship between fWHR and behavioral tendencies (and the possibility of an evolutionary mismatch) by explicitly addressing the above-mentioned limitations in prior research (i.e.,

the reliance on small sample sizes or self-reported data). Specifically, we conducted a large-scale study on the association between fWHR and behavioral tendencies that employed both self- and other-rated measures in a sample of business executives. An additional positive and original feature of our study is that we obtained our other-rated measures of behavioral tendencies from people who reported knowing the focal individual well, thus reducing the possibility that fWHR was simply biasing the judgment of raters who were unfamiliar with the person (e.g., Efferson & Vogt, 2013; Geniole et al., 2017). We see our study's use of a very large sample and other-reported ratings from people familiar with the target as providing the foundation for a clearer assessment of whether fWHR predicts antisocial tendencies or whether the influence of fWHR on social perceptions may be an example of an evolutionary mismatch.

## Method

### Sample

The source of our data was the flagship executive-development program of the Center for Creative Leadership (CCL). CCL is a nonprofit organization headquartered in Greensboro, North Carolina, that specializes in leadership training. The initial sample in this proprietary data set consisted of 1,305 executives who participated in a feedback survey also completed by the executives' colleagues called the Campbell Leadership Index (CLI; Campbell, 1991) from 2014 to 2017. We were able to capture not only self-reported ratings but also ratings from other people who reported knowing the focal executives well (subordinates, peers, and supervisors). To gather the self-reported data, CCL surveyed the focal executives at the start of the leadership program. For the other-reported data, colleagues from the focal executives' firms were surveyed anonymously prior to the start of the program.

After accounting for missing pictures of the executives ( $n = 126$ ) and averaging ratings within rater category for categories that included multiple raters per person (Atwater, Wang, Smither, & Fleenor, 2009), we had a final sample of 1,179 executives (873 males, 306 females) for whom we had a complete set of self-, peer, subordinate, and superior ratings. Interrater reliability was acceptable ( $\alpha = .72$  for peer ratings,  $\alpha = .61$  for subordinate ratings, and  $\alpha = .66$  for superior ratings).

### Estimating fWHR

We ensured that the executives were facing forward in all pictures. Some of the facial images were slightly

tilted to the right or left, which would have affected our fWHR measurements. To correct for this, we rotated these pictures so the eyes were on a horizontal plain. We then standardized the pictures to 8-bit gray-scale images with a height of 400 pixels (Carré, McCormick, & Mondloch, 2009). Next, two research assistants independently measured the facial width (distance between the right and left zygion) and facial height (distance between the midbrow and upper lip) of every executive using the National Institutes of Health's ImageJ software (Abràmoff, Magalhães, & Ram, 2004). We computed fWHR as facial width divided by facial height. Given that the reliability score was high ( $r = .90$  for fWHR;  $r = .94$  for both width and height), we averaged the scores between the two raters.

We sought to further validate the results by measuring fWHR using the Facial Attributes function of Face++, an online artificial intelligence (AI) application (Kosinski, 2017), and found strong reliability between human raters and AI raters ( $\alpha = .81$ ), giving us greater confidence in our fWHR measurements. The mean human-rated fWHR for males was 1.81, 95% confidence interval (CI) = [1.802, 1.818], and for females was 1.70, 95% CI = [1.687, 1.713], whereas the mean AI-rated fWHR for males was 1.90, 95% CI = [1.891, 1.909], and for females was 1.85, 95% CI = [1.837, 1.863].

### **Behavioral measures**

From the 100 items in the CLI survey, we chose the 25 that reflected anti- and prosocial or desirable and undesirable behavioral tendencies. These 25 items and their descriptions are reported in the Supplemental Material available online. To increase interpretability, we ran a common factor analysis to extract latent variables from the 25 items. First, the common factor analysis indicated that there were only three factors with eigenvalues greater than 1. A number of items cross-loaded onto multiple factors. We dropped items with loadings that were less than .40 and obtained the most optimal model of the three factors. On the basis of prior research (Fiske, Cuddy, Glick, & Xu, 2002; Goodwin, Piazza, & Rozin, 2014; Stavrova & Ehlebracht, 2016), we labeled these factors warmth (items: considerate, sensitive, affectionate, friendly, likeable, insensitive), cynicism (items: suspicious, cynical, temperamental, resentful, sarcastic), and morality (items: ethical, credible, candid, deceptive). The factor loadings are reported in Table 1. Table 2, which provides the correlations between self- and other-ratings on these three factors, shows that the correlation between different rater groups was low. This is consistent with our argument that self-reports and other-reports are likely to differ.

### **Results**

Table 3 shows the correlations between fWHR and the three factors from the CLI items (i.e., cynicism, morality, and warmth). We found that for males, both human-rated ( $p = .037$ ) and AI-rated ( $p = .016$ ) fWHR were positively related to self-rated cynicism but not to other-rated cynicism. Additionally, we found that AI-rated fWHR was negatively correlated with superior-rated cynicism for females ( $p = .048$ ). All other correlations were nonsignificant. Table 4 shows the partial correlations between fWHR and the three factors after we controlled for age, race, and whether the target individual was smiling in the picture. After including these controls, we found that only one significant relationship remained. Whereas human-rated fWHR continued to be positively related to self-rated cynicism for males ( $p = .048$ ), the positive relationship between AI-rated fWHR and self-rated cynicism for males became nonsignificant by a small margin ( $p = .052$ ). When we applied a more conservative test of significance that adjusted for multiple comparisons ( $p < .05$  divided by 24 within-gender comparisons), we failed to find any significant bivariate or partial correlations (see Tables S1 and S2 in the Supplemental Material).

As an additional test of whether there was a relationship between fWHR and behavioral tendencies, we computed Bayes factor ( $BF_{01}$ ) values for the bivariate and partial correlations (reported in Tables 3 and 4, respectively).  $BF_{01}$  measures the likelihood that the observed data fit the null hypothesis (i.e., there is no relationship between fWHR and behavioral tendencies) as opposed to the alternative hypothesis (i.e., there is a relationship between fWHR and behavioral tendencies), whereas  $BF_{10}$  ( $1/BF_{01}$ ) measures the likelihood that the data fit the alternative hypothesis as opposed to the null hypothesis (Aczel et al., 2018; Wagenmakers, Morey, & Lee, 2016). We used  $BF_{01}$  for our analysis since we were interested in the likelihood of the null hypothesis over the alternative hypothesis.  $BF_{01}$  values above 1 but below 3 provide anecdotal support for the null hypothesis,  $BF_{01}$  values above 3 but below 10 provide moderate support for the null hypothesis, and  $BF_{01}$  values above 10 provide strong support for the null hypothesis; the same is the case for  $1/BF_{01}$ , except that values in the given ranges provide support for the alternative rather than the null hypothesis. As the results in Tables 3 and 4 indicate, the  $BF_{01}$  values for almost all of the bivariate and partial correlations were above 1, supporting the null hypothesis of no relationship between fWHR and behavioral tendencies. The one exception was the bivariate correlation between AI-rated fWHR and self-reported cynicism for males ( $BF_{01} = 0.718$ ), indicating that it is 1.39 times more likely ( $1/0.718$ ) that the data

**Table 1.** Results of the Common Factor Analysis of Selected Campbell Leadership Index Items With Uniqueness Score and Percentage of Variance Explained

Item	Warmth	Cynicism	Morality	Uniqueness
Sensitive (highly aware of the feelings of others)	.756	—	—	.339
Considerate (thoughtful of the needs and feelings of others)	.754	—	—	.292
Affectionate (acts close, warm, and caring toward others)	.727	—	—	.424
Friendly (warm and pleasant, nice to be around)	.710	—	—	.265
Likeable (easy to feel friendly toward)	.669	—	—	.298
Insensitive (unaware of the feelings of others)	-.645	—	—	.393
Suspicious (inclined to distrust others)	—	.630	—	.441
Temperamental (moody, irritable, and overly sensitive)	—	.603	—	.493
Cynical (doubts the goodness of others)	—	.597	—	.494
Resentful (feels injured, insulted, or exploited)	—	.580	—	.559
Sarcastic (makes cutting remarks belittling others)	—	.570	—	.517
Credible (worthy of trust, believable)	—	—	.679	.372
Candid (open and honest when dealing with others)	—	—	.560	.539
Ethical (lives within society's standards of right and wrong)	—	—	.519	.632
Deceptive (conceals the truth for selfish reasons)	—	—	-.560	.493
Variance explained	50.87%	31.53%	24.23%	

Note: Only factor loadings above .40 are reported. Items were rotated using varimax rotation.

conform to the prediction that there is (vs. is not) a relationship between AI-rated fWHR and self-reported cynicism for males.

We further assessed the robustness of our results by conducting two supplementary analyses, the results of which are available in Tables S3 and S5 in the Supplemental Material. First, we conducted subsample analyses to assess whether fWHR would be more strongly related to the three factors when we examined individuals who had only extreme fWHR values, which we defined as those 361 executives whose fWHR values fell either 1 standard deviation above or 1 standard deviation below the mean values. The results from this supplementary subsample analysis suggested that our

previously discussed findings are robust: No significant correlations were obtained, with the exception that human-rated fWHR ( $p = .018$ ) and AI-rated fWHR ( $p = .004$ ) were positively correlated with self-rated cynicism among males, and the  $1/BF_{01}$  values for both of these correlations were above 1 (2.19 and 7.81, respectively). When we adjusted for multiple comparisons (see Table S4 in the Supplemental Material), these correlations became nonsignificant.

As a second robustness check, we also examined the correlations between fWHR and each of the original 25 CLI items and found no consistent patterns of significance. Furthermore, the significant relationships that were found became nonsignificant after we adjusted

**Table 2.** Correlations Between Self-, Peer-, Subordinate-, and Superior-Reported Scores on 3 Factors From the 25 Campbell Leadership Index Items ( $N = 1,179$ )

Factor and report type	Peer			Self			Subordinate		
	<i>r</i>	<i>p</i>	95% CI	<i>r</i>	<i>p</i>	95% CI	<i>r</i>	<i>p</i>	95% CI
Cynicism									
Self	.284	< .001	[.231, .336]						
Subordinate	.404	< .001	[.355, .451]	.264	< .001	[.210, .316]			
Superior	.370	< .001	[.319, .418]	.188	< .001	[.133, .243]	.279	< .001	[.226, .331]
Morality									
Self	.131	< .001	[.074, .187]						
Subordinate	.324	< .001	[.272, .374]	.197	< .001	[.141, .251]			
Superior	.303	< .001	[.250, .354]	.123	< .001	[.066, .179]	.253	< .001	[.199, .306]
Warmth									
Self	.424	< .001	[.376, .470]						
Subordinate	.494	< .001	[.449, .536]	.374	< .001	[.324, .422]			
Superior	.465	< .001	[.419, .508]	.379	< .001	[.329, .427]	.401	< .001	[.352, .448]

Note: CI = confidence interval.

**Table 3.** Correlations Between Facial Width-to-Height Ratio (fWHR) and Self-, Peer-, Subordinate-, and Superior-Reported Scores on 3 Factors From the 25 Campbell Leadership Index Items

Factor and report type	Males ( <i>n</i> = 873)				Females ( <i>n</i> = 306)			
	<i>r</i>	<i>p</i>	BF <sub>01</sub>	95% CI	<i>r</i>	<i>p</i>	BF <sub>01</sub>	95% CI
fWHR rater (human)								
Cynicism								
Peer	.029	.386	8.675	[−.037, .096]	.025	.659	6.816	[−.087, .137]
Self	.071	.037	1.441	[.004, .136]	−.046	.426	5.490	[−.157, .067]
Subordinate	−.008	.807	12.236	[−.075, .058]	−.073	.202	3.366	[−.184, .039]
Superior	.018	.602	11.011	[−.049, .084]	−.029	.612	6.609	[−.141, .083]
Morality								
Peer	−.018	.594	1.943	[−.084, .048]	.032	.575	6.427	[−.08, .144]
Self	−.002	.947	12.578	[−.069, .064]	−.011	.851	7.370	[−.123, .101]
Subordinate	.03	.378	8.565	[−.037, .096]	−.028	.624	6.665	[−.14, .084]
Superior	.047	.162	4.766	[−.019, .113]	.05	.386	5.186	[−.063, .161]
Warmth								
Peer	−.01	.763	12.048	[−.077, .056]	.025	.662	6.826	[−.087, .137]
Self	.01	.776	12.110	[−.057, .076]	.037	.524	6.145	[−.076, .148]
Subordinate	.01	.76	12.035	[−.056, .077]	.034	.554	6.311	[−.078, .146]
Superior	.021	.528	1.341	[−.045, .088]	−.014	.809	7.286	[−.126, .098]
fWHR rater (artificial intelligence)								
Cynicism								
Peer	−.023	.505	1.111	[−.089, .044]	−.01	.867	7.395	[−.122, .103]
Self	.081	.016	0.718	[.015, .147]	−.004	.94	7.477	[−.116, .108]
Subordinate	−.035	.305	7.481	[−.101, .032]	−.083	.148	2.689	[−.193, .03]
Superior	−.013	.711	11.773	[−.079, .054]	−.113	.048	1.092	[−.223, −.001]
Morality								
Peer	.037	.271	6.909	[−.029, .103]	.001	.992	7.498	[−.112, .113]
Self	.021	.544	1.497	[−.046, .087]	.058	.315	4.567	[−.055, .169]
Subordinate	.012	.723	11.844	[−.054, .078]	.03	.604	6.570	[−.083, .141]
Superior	.055	.103	3.358	[−.011, .121]	.064	.267	4.090	[−.049, .175]
Warmth								
Peer	.031	.362	8.345	[−.036, .097]	.052	.367	5.024	[−.061, .163]
Self	.003	.919	12.541	[−.063, .07]	.063	.274	4.168	[−.05, .174]
Subordinate	−.004	.909	12.523	[−.07, .062]	.102	.073	1.554	[−.01, .212]
Superior	.044	.192	5.407	[−.022, .11]	.089	.122	2.319	[−.024, .199]

Note: BF = Bayes factor; CI = confidence interval.

for multiple comparisons (see Table S6 in the Supplemental Material). Taken together, the lack of statistically detectable relationships between fWHR and behavioral tendencies (excepting fWHR and self-reported cynicism in males) leads us to conclude that the influence of fWHR on the social judgments of evaluators who are unfamiliar with targets (as documented in prior research) is likely a case of evolutionary mismatch.

## Discussion

We began by noting that what had appeared to be a growing consensus that fWHR predicts a host of anti-social tendencies (e.g., Anderl et al., 2016; Carré &

McCormick, 2008; Goetz et al., 2013; Haselhuhn & Wong, 2012; Stirrat & Perrett, 2010) had been challenged by Kosinski (2017), who notably found little evidence supporting this relationship. Our hope in this study is to have advanced this emerging debate by extending prior research, including Kosinski's study, in several important ways. Specifically, we sought to improve on prior fWHR research by conducting a large-scale study on the relationship between fWHR and behavioral tendencies that did not rely exclusively on self-reported data (and their attendant social-desirability bias). Instead, our survey data were provided by focal individuals and by evaluators, and these evaluators reported knowing the focal individuals well. With

**Table 4.** Partial Correlations Between Facial Width-to-Height Ratio (fWHR) and Self-, Peer-, Subordinate-, and Superior-Reported Scores on 3 Factors From the 25 Campbell Leadership Index Items, With Controls

Factor and report type	Males ( <i>n</i> = 873)				Females ( <i>n</i> = 306)			
	<i>r</i>	<i>p</i>	BF <sub>01</sub>	95% CI	<i>r</i>	<i>p</i>	BF <sub>01</sub>	95% CI
fWHR rater (human)								
Cynicism								
Peer	.046	.176	5.179	[−.021, .112]	.030	.603	6.579	[−.083, .142]
Self	.067	.048	1.742	[.001, .133]	−.036	.530	6.085	[−.148, .076]
Subordinate	.001	.974	12.599	[−.065, .067]	−.076	.186	3.385	[−.186, .037]
Superior	.029	.398	8.898	[−.038, .095]	−.021	.711	7.066	[−.133, .091]
Morality								
Peer	−.034	.321	7.731	[−.100, .033]	.020	.731	7.020	[−.093, .132]
Self	−.010	.774	12.080	[−.076, .057]	−.019	.738	7.188	[−.131, .093]
Subordinate	.028	.408	8.921	[−.038, .094]	−.031	.594	6.534	[−.142, .082]
Superior	.033	.330	7.842	[−.033, .099]	.045	.435	5.528	[−.068, .156]
Warmth								
Peer	−.020	.563	1.738	[−.086, .047]	.033	.568	6.461	[−.080, .144]
Self	.009	.788	12.183	[−.057, .075]	.039	.496	6.036	[−.073, .150]
Subordinate	.010	.769	12.054	[−.056, .076]	.036	.535	6.344	[−.077, .147]
Superior	.018	.596	1.961	[−.048, .084]	−.012	.832	7.305	[−.124, .100]
fWHR rater (artificial intelligence)								
Cynicism								
Peer	−.001	.984	12.602	[−.067, .066]	.022	.705	6.993	[−.091, .134]
Self	.066	.052	1.888	[−.001, .132]	.005	.925	7.468	[−.107, .117]
Subordinate	−.030	.374	8.512	[−.096, .036]	−.083	.147	2.748	[−.193, .029]
Superior	.004	.897	12.505	[−.062, .071]	−.092	.109	2.160	[−.202, .020]
Morality								
Peer	.027	.418	9.121	[−.039, .094]	−.011	.842	7.366	[−.123, .101]
Self	.014	.682	11.612	[−.053, .080]	.064	.265	4.023	[−.049, .175]
Subordinate	.023	.501	1.061	[−.044, .089]	.025	.669	6.852	[−.088, .136]
Superior	.038	.258	6.690	[−.028, .104]	.043	.449	5.642	[−.069, .155]
Warmth								
Peer	.017	.618	11.137	[−.050, .083]	.062	.283	4.273	[−.051, .172]
Self	.001	.984	12.604	[−.066, .067]	.057	.323	4.662	[−.056, .168]
Subordinate	−.006	.852	12.393	[−.073, .060]	.096	.095	1.954	[−.017, .206]
Superior	.039	.252	6.589	[−.028, .105]	.081	.156	2.797	[−.031, .192]

Note: Campbell Leadership Index factors and fWHR were obtained by regressing facial expression, age, and race and then obtaining the residuals. BF = Bayes factor; CI = confidence interval.

this improved research design, and after conducting a variety of analyses, our results clearly suggest that there is very little evidence in support of a relationship between fWHR and antisocial tendencies (excepting some evidence on self-rated cynicism in males).

We interpret our findings as more consistent with the notion of an evolutionary mismatch, whereby fWHR, once reliably tied to antisocial tendencies in ancestral environments in which violence was far more pervasive, may no longer be predictive of these tendencies in modern environments, leading to biased perceptions (Li et al., 2018). Indeed, we suggest that the fact that we found no reliable link between fWHR and

behavior does not necessarily mean that fWHR is unrelated to social judgments. Existing research has found that fWHR is positively related to a number of antisocial perceptions, including aggressiveness (Lefevre & Lewis, 2014), deceptiveness (Efferson & Vogt, 2013), proneness to anger (Deska et al., 2018a), threat potential (Geniole et al., 2017), and ascriptions of inhumanity (Deska et al., 2018b). We welcome future research that further refines our understanding of how and why physical characteristics, such as fWHR, may differentially affect focal-actor behaviors and social perceptions. Such an endeavor will increase our understanding of not only our evolved psyche but also why immutable physical

characteristics may shape social outcomes, even when they are not predictive of actual behavior.

While our findings appear robust, we also want to acknowledge two caveats. First, our results could have been affected by ceiling and reference-group effects because the business executives in our sample may differ systematically on both fWHR (Alrajih & Ward, 2014) and antisocial tendencies (Babiak, Neumann, & Hare, 2010) compared with the general population. Although the mean fWHR for males and females in our sample was similar to the general population (see Alrajih & Ward, 2014; Kosinski, 2017), we could not rule out the possibility that the executives in our sample are more antisocial (Babiak et al., 2010). Second, since our data were collected in the context of occupational training, it is possible that respondents could have been hesitant to portray their colleagues in a negative light. While we could not entirely rule out this possibility, the leadership training program attempted to reduce this bias by making the ratings completely anonymous and by making it clear to the raters that only scores aggregated across all raters would be disclosed to individual participants. We welcome future research that extends our work to consider these and other issues relevant to the growing body of evolutionary-psychology research linking physical characteristics and behavioral tendencies.

### Action Editor

D. Stephen Lindsay served as action editor for this article.

### Author Contributions

D. Wang, K. Nair, M. Kouchaki and E. J. Zajac developed the study concept. X. Zhao provided the data. D. Wang analyzed and interpreted the data under the supervision of K. Nair and M. Kouchaki. K. Nair and D. Wang drafted the manuscript under the guidance of M. Kouchaki and E. J. Zajac. E. J. Zajac provided critical revisions. All authors approved the final version of the manuscript for submission.

### Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

### Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797619849928>

### Open Practices

The design and analysis plans for the present study were not preregistered. Researchers interested in obtaining the data can contact the Center for Creative Leadership (<https://www.ccl.org/>) or X. Zhao directly (zhaos@ccl.org).

### References

- Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics international*, 11(7), 36–42.
- Aczel, B., Palfi, B., Szollosi, A., Kovacs, M., Szaszi, B., Szecsi, P., . . . Wagenmakers, E.-J. (2018). Quantifying support for the null hypothesis in psychology: An empirical investigation. *Advances in Methods and Practices in Psychological Science*, 1, 357–366.
- Alrajih, S., & Ward, J. (2014). Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *British Journal of Psychology*, 105, 153–161.
- Anderl, C., Hahn, T., Schmidt, A. K., Moldenhauer, H., Notebaert, K., Clément, C. C., & Windmann, S. (2016). Facial width-to-height ratio predicts psychopathic traits in males. *Personality and Individual Differences*, 88, 99–101.
- Atkins, P. W. B., & Wood, R. E. (2002). Self- versus others' ratings as predictors of assessment center ratings: Validation evidence for 360-degree feedback programs. *Personnel Psychology*, 55, 871–904.
- Atwater, L., Wang, M., Smither, J. W., & Fleenor, J. W. (2009). Are cultural characteristics associated with the relationship between self and others' ratings of leadership? *Journal of Applied Psychology*, 94, 876–886.
- Babiak, P., Neumann, C. S., & Hare, R. D. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences & the Law*, 28, 174–193.
- Bird, B. M., Cid Jofré, V. S., Geniole, S. N., Welker, K. M., Zilioli, S., Maestripieri, D., . . . Carré, J. M. (2016). Does the facial width-to-height ratio map onto variability in men's testosterone concentrations? *Evolution and Human Behavior*, 37, 392–398.
- Campbell, D. (1991). *Campbell Leadership Index (CLI)*. Minneapolis, MN: National Computer Systems.
- Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London B: Biological Sciences*, 275, 2651–2656.
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, 20, 1194–1198.
- Daly, M., & Wilson, M. (1988). *Homicide*. Piscataway, NJ: Transaction Publishers.
- Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018a). The face of fear and anger: Facial width-to-height ratio biases recognition of angry and fearful expressions. *Emotion*, 18, 453–464.
- Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018b). Facing humanness: Facial width-to-height ratio predicts ascriptions of humanity. *Journal of Personality and Social Psychology*, 114, 75–94.
- Efferson, C., & Vogt, S. (2013). Viewing men's faces does not lead to accurate predictions of trustworthiness. *Scientific Reports*, 3, Article 1047. doi:10.1038/srep01047
- Eisenbruch, A. B., Lukaszewski, A. W., Simmons, Z. L., Arai, S., & Roney, J. R. (2018). Why the wide face? Androgen

- receptor gene polymorphism does not predict men's facial width-to-height ratio. *Adaptive Human Behavior and Physiology*, 4, 138–151.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878–902.
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLOS ONE*, 10(7), Article e0132726. doi:10.1371/journal.pone.0132726
- Geniole, S. N., Keyes, A. E., Carré, J. M., & McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences*, 57, 59–64.
- Geniole, S. N., MacDonell, E. T., & McCormick, C. M. (2017). The threat premium in economic bargaining. *Evolution and Human Behavior*, 38, 572–582.
- Goetz, S. M., Shattuck, K. S., Miller, R. M., Campbell, J. A., Lozoya, E., Weisfeld, G. E., & Carré, J. M. (2013). Social status moderates the relationship between facial structure and aggression. *Psychological Science*, 24, 2329–2334.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 148–168.
- Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society of London B: Biological Sciences*, 279, 571–576.
- Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science*, 28, 1675–1682.
- Lefevre, C. E., & Lewis, G. J. (2014). Perceiving aggression from facial structure: Further evidence for a positive association with facial width-to-height ratio and masculinity, but not for moderation by self-reported dominance. *European Journal of Personality*, 28, 530–537.
- Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, 34, 273–279.
- Li, N. P., van Vugt, M., & Colarelli, S. M. (2018). The evolutionary mismatch hypothesis: Implications for psychological science. *Current Directions in Psychological Science*, 27, 38–44.
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, 69, 511–534.
- Stavrova, O., & Ehlebracht, D. (2016). Cynical beliefs about human nature and income: Longitudinal and cross-cultural analyses. *Journal of Personality and Social Psychology*, 110, 116–132.
- Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, 21, 349–354.
- Stirrat, M., Stulp, G., & Pollet, T. V. (2012). Male facial width is associated with death by contact violence: Narrow-faced males are more likely to die from contact violence. *Evolution and Human Behavior*, 33, 551–556.
- Wagenmakers, E.-J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, 25, 169–176.