Personality and Individual Differences 174 (2021) 110683

Contents lists available at ScienceDirect

Personality and Individual Differences

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

Facial width-to-height ratios and deception skill

David Matsumoto [a,b,*], Hyisung C. Hwang [b]

[a] San Francisco State University, United States of America
[b] Humintell, PO Box 1304, El Cerrito, CA 94530, United States of America

ARTICLE INFO

Keywords:
Facial width-to-height ratio
Deception
Ethnicity

ABSTRACT

Prior research on facial width-to-height ratio (fWHR) has demonstrated links between it and numerous traits and antisocial behavior, including deception. No study, however, has examined whether fWHR is associated with deception skills. Here, a community sample of $N = 72$ individuals were randomly assigned to truth or lie conditions in a mock theft experiment and interviewed. Video clips from the interviews were shown to $N = 248$ observers, who judged whether the interviewees were telling the truth or lying, the likelihood that interviewees committed a theft, and confidence in their judgments. There was no association between fWHR and accuracy scores or confidence ratings, and good and bad truthtellers and liars did not differ in their fWHR. Instead, people with larger fWHR, regardless of veracity condition, sex, or ethnicity, were more likely to be judged inaccurately as lying and as having committed the crime. fWHR may influence social perceptions about deception.

1. Introduction

Facial width-to-height ratio (fWHR) has increasingly received attention as an indicator of genetic fitness, with prior research demonstrating links between fWHR and numerous traits and behavior. For example, fWHR has been associated with dominance and aggression (Lefevre et al., 2014); perceptions of intimidation (Hehman, Leitner, & Gaertner, 2013); violent or antisocial behavior (Carre & McCormick, 2008); judgments of angry and fearful facial expressions (Deska, Paige, & Hugenberg, 2018) and humanness (Deska, Lloyd, & Hugenberg, 2018); perceived dominance, likelihood of being chosen for a second date, and attractiveness to women for short-term relationships (Mileva et al., 2014; Valentine et al., 2014); testosterone levels (Lefevre et al., 2013); deception and cheating (Haselhuhn & Wong, 2012); trust and trustworthiness (Stirrat & Perrett, 2010); prejudicial beliefs (Hehman, Leitner, Deegan, & Gaertner, 2013); sexual desire (Arnocky et al., 2018); organizational success (Wong et al., 2011), and achievement drive in US presidents (Lewis et al., 2012). A meta-analysis summarizing this area of research indicated that fWHR predicted threat behavior in males and dominance behavior in both sexes, and that individuals with larger fWHRs were judged as more threatening and dominant and less attractive than those with smaller fWHRs, suggesting that fWHR is an evolved cue of threat (Geniole et al., 2015).

A few studies have also raised questions about some of these links, demonstrating that fWHR was not associated with testosterone during adolescence (Hodges-Simeon et al., 2016) or with sexual desire in women (Zhang et al., 2018). A study examining the association between fWHR and behavioral tendencies in 137,163 participants and 55 scales of intelligence, personality, impulsiveness, sense of fairness, sensational interests, self-monitoring, impression management, and satisfaction with life demonstrated that fWHR was not meaningfully linked with any of these self-reported measures (Kosinski, 2017). Another study of more than 1000 business executives involving not only self-reports but also third-party ratings found little evidence of fWHR predicting antisocial tendencies (Wang et al., 2019). These authors suggested that fWHR may be a physical characteristic that once was associated with antisocial tendencies in evolutionary history but is not predictive of such behavior today.

An unexplored area in this genre of research concerns whether fWHR is associated with skills in engaging in antisocial behavior, notably deception. Although one study demonstrated that male fWHR was positively associated with deceptive behavior in a negotiation exercise and cheating in a dice game (Haselhuhn & Wong, 2012), no study has addressed whether fWHR is associated with deception skill. Such an examination would require individuals with varying fWHR not only to engage in truthful or deceptive behavior, but then to have naïve observers judge the individuals as truthful or deceptive. Individuals who engaged in deception and fooled others (i.e., were judged to be truthful) could be considered skillful at deception. Those who engaged in deception and were judged to be deceptive would be less deceptively

* Corresponding author at: Department of Psychology, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, United States of America.
E-mail addresses: dm@sfsu.edu (D. Matsumoto), hhwg51@gmail.com (H.C. Hwang).

https://doi.org/10.1016/j.paid.2021.110683
Received 12 November 2020; Received in revised form 18 January 2021; Accepted 19 January 2021
Available online 30 January 2021

skillful.

Examining associations between fWHR and deceptive skills is related to literature on individual differences in deception. Although previous studies have examined individual differences on deceptive behavior (Vrij et al., 1997) and deception judgments (Bond & DePaulo, 2008), only a few have examined individual differences in deceptive skills. These studies have focused on psychological traits including dominance, extraversion, exhibition, expressiveness, emotional sending, nurturance, and play (Riggio & Friedman, 1983); sender demeanor (Levine et al., 2011); and outgoingness, energy, apprehensiveness, and social desirability (Riggio et al., 1988); one study examined age (Ruffman et al., 2012). To date, no study has examined fWHR and deception skills.

The purpose of this study was to address this issue. In Phase 1, a community sample of individuals of two ethnicities participated in a mock theft experiment, were randomly assigned to commit a theft or not, and were interviewed about the theft at different times during an experiment. In Phase 2, videoclips of their interviews were shown to naïve groups of observers who judged whether or not interviewees were truthful, and rated the likelihood that they actually committed the theft and confidence in their judgments. Because ground truth was known, accuracy scores of the observers' judgments were computed. We hypothesized that fWHR would be positively associated with greater deception skills, i.e., less accuracy in deception judgments, ratings of the likelihood of having committed the theft, and confidence ratings.

2. Method

2.1. Phase 1: stimuli production

A convenience sample of $N = 72$ individuals recruited from the community and comprised of $n = 34$ Chinese (20 females) and $n = 38$ European Americans (18 females) participated in a mock crime experiment and were randomly assigned to truth ($n = 34$) or lie ($n = 38$) conditions (details about the methodology of the stimuli production can be found in (Matsumoto & Hwang, 2018; Matsumoto & Hwang, 2020). The two groups were originally included in order to examine cultural/ethnic differences in truthful and deceptive behavior, and their inclusion in the current study allowed for examination of generalization of findings across ethnicities in associations of fWHR and deceptive skill. The mock theft feature of the study allowed participants a chance to steal money or not, and then to tell the truth or lie about their actions to interviewers. Stakes were associated with being believed or not by the interviewers (possibilities of more or less compensation and requirement to stay longer or not). Manipulation checks were assessed by self-reported emotions before and after the interviews and by direct ratings of the perceived stakes, and confirmed that participants were emotionally aroused differentially by veracity condition and perceived moderate-high stakes associated with their performances.

The experiment included three interviews, two before the chance to steal and one after. Each interview included a segment including questions to which interviewees had to lie or tell the truth directly about the theft; video clips were excerpted from these segments. Each segment was examined for protocol compliance and video clips were selected to avoid judgments based on procedural idiosyncrasies within interviews, including interviews in which responses revealed condition, interviewers' questioning was atypical, and segments were of similar duration. This resulted in a final section of 109 video clips (mean duration = 97 s).

2.2. Phase 2: observer judgments

The video clips were judged in groups by a total $N = 248$ university undergraduates participating in partial fulfillment of course requirements, who each observed 20 video clips (mean n judging each video clip = 47.24). Video clips were randomly assigned to groups with the caveat that no group judged the same interviewee more than once. For each video clip, observers made three judgments: whether the interviewee was telling the truth or lying; the likelihood that the interviewee committed the theft on a 7-point scale anchored 1, Not at all Likely to 7, Extremely Likely; and the confidence they had in their judgments between 0 and 100%.

Because ground truth was known from the experiment, accuracy rates for each video were computed. Data for each video clip were merged across all observer judgment data available for that video; in cases in which multiple video clips of the same interviewee were judged, data were averaged (weighted) across video clips of the interviewee so that each interviewee contributed only one data point to the analyses below. These procedures resulted in data for means for truth and lie call rates (number of times interviewees were judged truthful or lying), accuracy (number of correct judgments given known ground truth concerning interviewees' conditions), likelihood of committing the theft, and confidence ratings for each interviewee. Note that because the raw judgment data involved binary choices of truth vs. lie, the mean truth and lie call rates were mirror images of each other. The units of analysis for the analyses below were the $N = 72$ interviewees.

2.3. Facial width-to-height (fWHR) measurement

fWHR measurement followed the same procedures as described in previous studies (Carre & McCormick, 2008; Deska, Lloyd, & Hugenberg, 2018; Deska, Paige, & Hugenberg, 2018; Wang et al., 2019). For each interviewee, video frames with the best frontal image of the interviewees' faces were found and recorded. Each face was then imported into Adobe Photoshop that allowed 1715 × 945 pixels and 300 ppi resolution. Lines were drawn to demarcate distances between the top of the lip and lower part of the brow (facial height), and distances between the most lateral points of the face by the ears (bizygomatic width); the image ruler was used to record measurements. To establish reliability, these procedures were redone by a second coder on 40 cases. Reliability between the two assessments was high, $r(39) = 0.98$, $p < .001$. fWHR was then computed by dividing width by height measurements.

3. Results

3.1. Preliminary analyses

We computed descriptive statistics for all variables separately for the total group, liars and truthtellers, females and males, and Asian and Caucasian interviewees (Table 1). We then computed Veracity Condition (2) × Sex (2) × Ethnicity (2) overall ANOVAs on lie call rates; mean accuracy, likelihood of theft, confidence ratings; and fWHR. (Because truth call rates were mathematically related to lie call rates, these analyses were not performed.) There were no significant effects on lie call rates or likelihood of theft. The Veracity Condition main effect was significant on accuracy, $F(1, 64) = 4.51$, $p = .038$, $\eta_p^2 = 0.066$; truthful videos were more accurately detected than lie videos, but this effect appeared to be due to higher truth than lie call rates, reflecting a truth bias (Bond & DePaulo, 2006). The Veracity Condition main effect was also significant for fWHR, $F(1, 61) = 6.33$, $p = .014$, $\eta_p^2 = 0.094$; individuals in the lie condition had significantly larger fWHR than those in the truth condition (even though interviewees were randomly assigned to condition). The Ethnicity main effect was significant on confidence ratings, $F(1, 64) = 4.84$, $p = .031$, $\eta_p^2 = 0.070$; Caucasians received higher confidence ratings than Asians. No interactions were significant.

3.2. Main analyses

We computed Pearson rs with 1000 bootstraps between fWHR and lie call rates, mean accuracy, likelihood of theft, and confidence ratings (Table 2). fWHR was *not* correlated with mean accuracy or confidence ratings, but was correlated with lie call rates and likelihood of theft ratings (with confidence intervals not overlapping with zero). Thus,

Table 1
Descriptive statistics (Ms and SDs) on truth and lie call rates; means of accuracy, likelihood of theft, and confidence ratings; and fWHR.

	Truth call rate	Lie call rate	Accuracy	Likelihood of theft	Confidence	fWHR
Total group	0.54	0.46	0.50	3.69	75.63	1.86
	(0.16)	(0.16)	(0.17)	(0.56)	(2.33)	(0.17)
Liars	0.54	0.46	0.46	3.71	75.61	1.91
	(0.17)	(0.17)	(0.17)	(0.59)	(2.04)	(0.18)
Truthtellers	0.54	0.46	0.54	3.67	75.72	1.82
	(0.16)	(0.16)	(0.16)	(0.54)	(2.68)	(0.14)
Female interviewees	0.57	0.43	0.51	3.57	75.53	1.84
	(0.16)	(0.16)	(0.18)	(0.59)	(2.48)	(0.16)
Male interviewees	0.51	0.49	0.50	3.82	75.81	1.90
	(0.16)	(0.15)	(0.16)	(0.50)	(2.22)	(0.17)
Asian interviewees	0.56	0.44	0.50	3.61	75.02	1.85
	(0.14)	(0.14)	(0.16)	(0.47)	(2.23)	(0.16)
Caucasian interviewees	0.52	0.48	0.50	3.76	76.23	1.88
	(0.18)	(0.18)	(0.18)	(0.63)	(2.33)	(0.17)

Table 2
Pearson rs between fWHR and truth and lie call rates, and means of accuracy, likelihood of theft, and confidence ratings with 1000 bootstraps.

	Truth call rates	Lie call rates	Accuracy	Likelihood of theft	Confidence
r	−0.28*	0.28*	−0.12	0.28*	−0.03
LLCI	−0.50	0.04	−0.36	0.05	−0.23
ULCI	−0.04	0.50	0.11	0.50	0.16

* $p < .05$; LLCI – Lower Level Confidence Interval; ULCI – Upper Level Confidence Interval.

observers were more likely to judge interviewees with larger fWHR as lying and having stolen the check, but were not necessarily accurate or confident in doing so.

To examine whether associations between fWHR and lie call rates or likelihood of theft ratings were moderated by either Veracity Condition, Sex, or Ethnicity, we computed separate moderation analyses using SPSS Process (Hayes, 2013), Model 1. No effect involving Condition, Sex, or Ethnicity was significant.

Because veracity condition, ethnicity, or sex of the interviewees may have covaried with the correlations reported in Table 2 (note the significant main effect of Veracity condition on fWHR reported above), we computed a hierarchical regression entering Veracity Condition, Ethnicity, and Sex as covariates on the first step (forced entry) and fWHR on the second, with lie call rates as the dependent. The first step was not significant, $R(64) = 0.21$, $p = .394$, but the second step was, $R^2_{chg}(63) = 0.07$, $p = .029$, as was the standardized coefficient for fWHR, $\beta = 0.28$, $t(63) = 2.24$, $p = .029$, CI [0.04, 0.57]. Thus, fWHR was associated with lie call rates even when veracity condition, ethnicity, and sex were controlled.

We also analyzed the data separately by good vs. bad truthtellers and liars. Interviewees in the truth condition whose videos received accuracy score means >0.50 were labeled good truthtellers, while those with accuracy score means <0.50 were labeled as bad truthtellers. Similarly, interviewees in the lie condition whose videos received accuracy scores >0.50 and < 0.50 were labeled as bad and good liars, respectively. We computed *t*-tests comparing good and bad truthtellers and liars on fWHR. Neither was significant, $t(31) = 0.73$, $p = .473$, $d = -0.25$; $t(34) = 0.70$, $p = .492$, $d = 0.25$, respectively. Thus, the correlations reported in Table 2 were not affected by whether interviewees were good or bad at truthtelling or lying.

3.3. Post-hoc (to deal with truth bias in responses)

Higher truth call rates corresponded to a truth bias in judging deception and may have affected accuracy. To examine whether the findings would survive if differences in call rates were controlled, we went back to the original judgment data and computed unbiased hit rates (Wagner, 1993). (Because video clips were randomly distributed across different groups of observers, we were unable to compute these separately for interviewee sex or ethnicities or for each individual interviewee video, which would have allowed us to merge the unbiased hit rate data with the fWHR data for analyses.) For each observer, we computed unbiased hit rates separately for truth and lie videos, arcsine transformed the two unbiased hit rates, and then computed paired *t*-tests comparing accuracy rates for truths and lies, with 1000 bootstraps. Truthful videos ($M = 0.29$, $SD = 0.13$) were judged more accurately than lie videos ($M = 0.26$, $SD = 0.14$), $t(248) = 4.83$, $p < .001$, $d = 0.27$, $M_{bootstrap} = 0.04$, 95% CI [0.02, 0.05] even after differential call rates were controlled.

4. Discussion

Contrary to prediction, fWHR was not associated with deceptive skills, as there was no association between fWHR and accuracy scores or confidence ratings, and good and bad truthtellers and liars did not differ significantly in their fWHR. Instead, people with larger fWHR, regardless of veracity condition, sex, or ethnicity, were more likely to be inaccurately judged as lying and as having committed the crime. These latter findings were somewhat remarkable, given that interviewees were randomly assigned to veracity conditions. Thus, our data suggested that fWHR drives social perceptions about deception, but were not really related to actual behavior.

These findings suggested a nuanced perspective on understanding the association between fWHR and deceptive or trust-related behavior. Although fWHR may indeed be linked to various antisocial behavior related to threat and dominance, especially in men, some social influence effects of fWHR may occur because of biased social perceptions by others. This notion would be consistent with a previous meta-analysis that demonstrated that individuals with larger fWHR are judged as more threatening and dominant, and less attractive (Geniole et al., 2015). Our data also lend credence to the notion that fWHR may have been associated with antisocial tendencies in evolutionary history, but today is linked with biased perceptions (Wang et al., 2019).

This interpretation suggests a Pygmalion effect possibility to consider. Social judgments related to deception or other kinds of antisocial behavior of a target individual with relatively wider fWHR may influence observers' own behavior toward the target individual, which in turn may influence the target individual to react in antisocial ways. To our knowledge, no study has examined this possible interactive effect of social judgments of an individual on that same individual's behavior vis-à-vis fWHR, and the current findings open the door to such examinations.

Biased perceptions about fWHR and deceptive behavior may have implications to real-world phenomena, including situations involving credibility assessments and judgments of trustworthiness, especially in first impressions. Hiring interviews, financial decisions, business

determinations, law enforcement interactions, dating and romance, and everyday life involve making such decisions. Possible interactive effects of social judgments afforded by fWHR on self- and other-behavior may have consequences to these real-life situations as well. For example, when interacting with a stranger in a novel social context in which only limited information is available, first impressions may be a major contributor to quick decisions. Thus, the reported findings may be helpful for people who deal with others in dynamic social contexts for brief times to reduce biased social judgments.

This study was not conducted without limitations, including the fact that fWHR was not pre-assessed and experimentally manipulated into veracity conditions, concerns about the ecological validity of the judgment task for university undergrads, and the relatively small sample size affecting power. For example, although tests of differences between good and bad truthtellers and liars on fWHR were not statistically significant, the Cohen's *d* effect sizes (0.25) suggested the existence of small effects that were consistent with prediction (i.e., good liars had larger fWHR). Future studies with larger sample sizes and more power may be required to tease out these effects, and other social judgments should be examined vis-à-vis fWHR in various social contexts as well.

Author contributions

Both authors developed the study concept and contributed equally to the study design. HH oversaw the data collections and data cleaning. DM conducted the preliminary data analyses and wrote the first draft of the paper. Both authors contributed equally to editing the paper to final version and approved it for submission.

Research disclosure statement

All dependent variables or measures that were analyzed for this article's target research question have been reported in the Method section.

All levels of all independent variables or all predictors or manipulations, whether successful or failed, have been reported in the Method section.

The total number of excluded observations and the reasons for making those exclusions (if any) have been reported in the Method section.

Open practices statement

The study reported in this article was not formally preregistered. Neither the data nor the materials have been made available on a permanent third-party archive; requests for the data or materials can be sent via email to the lead author at dm@sfsu.edu.

Ethical considerations

Procedures for conducting the mock theft experiment were approved by the University at Buffalo, State University of New York Social and Behavioral Sciences Institutional Review Board. Written consent was obtained from all interviewees appearing in the video records after they were fully informed about the experimental procedures but before the interviews and once again after the experiment was completed. Procedures for collecting the judgment data by the observers were approved by the Wright State University Institutional Review Board.

References

Arnocky, S., Carré, J. M., Bird, B. M., Moreau, B. J. P., Vaillancourt, T., Ortiz, T., & Marley, N. (2018). The facial width-to-height ratio predicts sex drive, sociosexuality, and intended infidelity. *Archives of Sexual Behavior, 47*(5), 1375–1385. https://doi.org/10.1007/s10508-017-1070-x.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234. https://doi.org/10.1207/s15327957pspr1003_2.

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*(4), 477–492.

Carre, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggresive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B, 275*, 2651–2656. https://doi.org/10.1098/rspb.2008.0873.

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018). Facing humanness: Facial width-to-height ratio predicts ascriptions of humanity. *Journal of Personality and Social Psychology, 114*(1), 75–94. https://doi.org/10.1037/pspi0000110.

Deska, J. C., Paige, L. E., & Hugenberg, K. (2018). The face of fear and anger: Facial width-to-height ratio biases recognition of angry and fearful expressions. *Emotion, 18*(3), 453–464. https://doi.org/10.1037/emo0000328.

Geniole, S. N., Denson, T. F., Dixson, 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. https://doi.org/10.1371/journal.pone.0132726.

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences, 279*(1728), 571–576. https://doi.org/10.1098/rspb.2011.1193.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression based approach*. New York: Guilford Press.

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2013). Facial structure is indicative of explicit support for prejudicial beliefs. *Psychological Science, 24*(3), 289–296. https://doi.org/10.1177/0956797612451467.

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology, 49*(4), 747–754. https://doi.org/10.1016/j.jesp.2013.02.015.

Hodges-Simeon, C. R., Hanson Sobraske, K. N., Samore, T., Gurven, M., & Gaulin, S. J. C. (2016). Facial width-to-height ratio (fWHR) is not associated with adolescent testosterone levels. *PLoS One, 11*(4), Article e0153083. https://doi.org/10.1371/journal.pone.0153083.

Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science, 28*(11), 1675–1682. https://doi.org/10.1177/0956797617716929.

Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P., & Penton-Voak, I. S. (2014). Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biology Letters, 10*(10), Article 20140729. https://doi.org/10.1098/rsbl.2014.0729.

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*(4), 273–279. https://doi.org/10.1016/j.evolhumbehav.2013.03.005.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., … Lee, J. H. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research, 37*(3), 377–403. https://doi.org/10.1111/j.1468-2958.2011.01407.x.

Lewis, G. J., Lefevre, C. E., & Bates, T. C. (2012). Facial width-to-height ratio predicts achievement drive in US presidents. *Personality and Individual Differences, 52*(7), 855–857. https://doi.org/10.1016/j.paid.2011.12.030.

Matsumoto, David, & Hwang, Hyisung C. (2018). Clusters of nonverbal behaviors differ according to type of question and veracity in investigative interviews in a mock crime context. *Journal of Police and Criminal Psychology, 33*(4), 302–3105. https://doi.org/10.1007/s11896-017-9250-0.

Matsumoto, David, & Hwang, Hyisung C. (2020). Clusters of nonverbal behavior differentiate truths and lies about future malicious intent in checkpoint screening interviews. *Psychiatry, Psychology and Law*. https://doi.org/10.1080/13218719.2020.1794999.

Mileva, V. R., Cowan, M. L., Cobey, K. D., Knowles, K. K., & Little, A. C. (2014). In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences, 69*, 115–118. https://doi.org/10.1016/j.paid.2014.05.019.

Riggio, R. E., & Friedman, H. S. (1983). Individual differences and cues to deception. *Journal of Personality and Social Psychology, 45*(4), 899–915. https://doi.org/10.1037/0022-3514.45.4.899.

Riggio, R. E., Salinas, C., & Tucker, J. (1988). Personality and deception ability. *Personality and Individual Differences, 9*(1), 189–191. https://doi.org/10.1016/0191-8869(88)90050-5.

Ruffman, T., Murray, J., Halberstadt, J., & Vater, T. (2012). Age-related differences in deception. *Psychology and Aging, 27*(3), 543–549. https://doi.org/10.1037/a0023380.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*(3), 349–354. https://doi.org/10.1177/0956797610362647.

Valentine, K. A., Li, N. P., Penke, L., & Perrett, D. I. (2014). Judging a man by the width of his face: The role of facial ratios and dominance in mate choice at speed-dating events. *Psychological Science, 25*(3), 806–811. https://doi.org/10.1177/0956797613511823.

Vrij, A., Akehurst, L., & Morris, P. (1997). Individual differences in hand movements during deception. *Journal of Nonverbal Behavior, 21*(2), 87–102.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*, 3–28.

Wang, D., Nair, K., Kouchaki, M., Zajac, E. J., & Zhao, X. (2019). A case of evolutionary mismatch? Why facial width-to-height ratio may not predict behavioral tendencies.

Psychological Science, 30(7), 1074–1081. https://doi.org/10.1177/0956797619849928.

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A face only an investor could love: CEOs' facial structure predicts their firms' financial performance. *Psychological Science, 22*(12), 1478–1483. https://doi.org/10.1177/0956797611418838.

Zhang, W., Hahn, A. C., Cai, Z., Lee, A. J., Holzleitner, I. J., DeBruine, L. M., & Jones, B. C. (2018). No evidence that facial width-to-height ratio (fWHR) is associated with women's sexual desire. *PLoS One, 13*(7), Article e0200308. https://doi.org/10.1371/journal.pone.0200308.