Journal of Research in Personality 46 (2012) 235–238

Contents lists available at SciVerse ScienceDirect

Journal of Research in Personality

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

Brief Report

Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players

Robert O. Deaner *, Stefan M.M. Goetz, Kraig Shattuck, Tony Schnotala

Department of Psychology, Grand Valley State University, Allendale, MI 49401, USA

ARTICLE INFO

Article history:
Available online 28 January 2012

Keywords:
Aggression
Aggressiveness
Body size
Face perception
Fighting
Judgment accuracy
Overgeneralization
Width-to-height ratio
Sports

ABSTRACT

Studies indicate that facial characteristics may predict behavior, but it is unclear if this will hold within highly selective populations. One relevant characteristic is the face's width-to-height ratio (FWHR), a sexually dimorphic trait that has been shown to predict aggression. That FWHR may predict aggression within highly selective populations was suggested by Carré and McCormick's (2008) finding that professional hockey players with greater FWHRs accrued more penalties. We attempted to replicate this result using all NHL players. We also explored fighting penalties as another aggression measure and height and weight as additional aggression predictors. We found that body weight predicted substantial variance in aggression but FWHR did not. Thus, in highly selective populations, inferences based on faces may be inaccurate.

1. Introduction

Humans are famously prone to draw inferences about behavioral dispositions from facial characteristics. Such inferences are traditionally lamented as undesirable and inaccurate, yet there is mounting evidence that they frequently contain "a kernel of truth." For instance, observers' ratings of static, non-expressive faces have been shown to significantly predict strength (Sell et al., 2009), agreeableness, and extraversion (Kramer & Ward, 2010).

Nonetheless, it remains unclear to what extent facial characteristics can bolster decision making in ecologically relevant contexts. One reason to be skeptical that they will is that, although the relation between a facial characteristic and a behavior might hold within a general population or sub-population, it might not hold within the range that is ecologically relevant to an observer, especially within a highly selective population (Zebrowitz & Montepare, 2008). For example, Zebrowitz and Rhodes (2004) showed that participants' ratings of perceived intelligence predicted measured intelligence but only for individuals below the median in attractiveness; they found similar results for perceptions of health. Thus, if an individual was considering only those who were high in attractiveness, their inferences regarding intelligence and health would be inaccurate overgeneralizations (Zebrowitz & Montepare, 2008).

One objective characteristic that might be broadly predictive of behavior is the upper face's width-to-height ratio (hereafter FWHR). FWHR is a sexually dimorphic facial trait that emerges around puberty and is thought to reflect selection for inter-male competition (Weston, Friday, & Lio, 2007). Supporting this last idea are several studies associating variation in FWHR with aggression: Carré and McCormick (2008) reported that it predicted ice hockey penalties and reactive aggression in the lab; Stirrat and Perrett (2010) showed that it predicted male willingness to exploit others in a trust game; Haselhuhn and Wong (2011) showed that it predicted willingness to cheat for financial gain; Carré, McCormick, and Mondloch (2009) and Carré, Morrissey, Mondloch, and McCormick (2010) demonstrated that FWHR predicted perceptions of male aggressiveness; and Stirrat and Perrett (2010) showed that participants were less trusting of men with greater FWHRs and that manipulating this feature decreased trust. In addition, Christiansen and Winkler (1992) reported that, among the !Kung San of Namibia, men with a more violent history had greater bizygomatic breadths, the width component of FWHR.

Carré and McCormick's (2008; study 3) finding that FWHR predicted penalty minutes among NHL (i.e. professional) hockey players is apparently the strongest evidence yet that a facial characteristic predicts a behavioral disposition even in a highly selective population. Here we test the robustness of this relationship with data from all 30 NHL teams, rather than only the Canadian ones (Carré & McCormick, 2008).

We also explore if FWHR predicts major fighting penalties, which comprise about 11% of overall penalty minutes (data from nhl.com; hockeyfights.com). Fighting should be a superior measure of aggression because overall penalties include infractions which

* Corresponding author. Fax: +1 616 331 2480.
E-mail address: deanerr@gvsu.edu (R.O. Deaner).

0092-6566/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jrp.2012.01.005

do not involve aggression, such as illegal substitution and delay of game. In addition, it is plausible that there could be referee bias in assessing some kinds of penalties against players that might "look the part" of an aggressor. However, this seems extremely unlikely for fighting because fighting penalties in the NHL are assessed for unambiguous altercations where the individual(s) clearly attempt to inflict damage (Bernstein, 2006).

Finally, we investigate body weight and height as predictors of aggression. Although many previous studies have indicated an association between body size and aggression (e.g., Ishikawa, Raine, Lencz, Bihrle, & LaCasse, 2001), none apparently have investigated it within a highly selective population or compared the predictiveness of size relative to FWHR.

2. Materials and methods

We obtained information on and photographs of NHL players from the Sports Illustrated website (sportsillustrated.cnn.com) between 12/16/2011 and 1/3/2012. We obtained information on major fighting penalties from the Hockeyfights website (hockeyfights.com). We initially included all players for whom there was a high quality frontal photograph available. Following Carré and McCormick (2008), however, we excluded goalkeepers. The final sample included 520 players. For penalties, we divided each player's total career penalty minutes by their total number of career games. Using career totals, rather than season totals (Carré & McCormick, 2008), should reduce random error. Information on total career fights was unavailable so we used the number of fights during the 2010–2011 season. Because some players on rosters during the 2011–2012 season would not have played during the 2010–2011 season, for all analyses involving fights, we only included players who had played at least 60 career games, which generally would indicate those who had played substantially during the 2010–2011 season.

Following Carré and McCormick (2008), we measured faces using ImageJ (NIH open-source software); face height was measured from upper lip to brow; face width was measured from the left to the right zygion (bizygomatic width). Some photographs showed a frontally oriented face that was not oriented perfectly vertical; in these cases we rotated the face prior to measurement. Initial measurement training included replicating Carré & McCormick's, 2008 result ($r = .30$) for six Canadian teams from the 2007 to 2008 season, and we succeeded ($r = .29$). For players in the current study, inter-rater reliability for FWHR was high ($r = .90$). All analyses were conducted with Statistica 6.1 (Statsoft Inc., Tulsa, OK USA).

3. Results

There was a positive correlation between FWHR and penalties per game, but it did not reach statistical significance ($r(518) = .084$, $p = .057$). A difference test indicated that this correlation departed significantly ($p = .03$) from the one ($r(110) = .30$, $p = .005$) reported by Carré and McCormick (2008). Next, following Carré and McCormick, we computed the correlations separately for each team. Nineteen of the correlations were positive, eleven were negative, and three reached statistical significance, two of which were negative. Fig. 1 shows the frequency of team correlation values, as well as those reported in Carré and McCormick. The team correlations obtained here differed significantly from Carré and McCormick's ($t(34) = 2.13$, $p = .04$).

Similar to Carré and McCormick (2008), we found that Canadian teams showed predominantly positive correlations (Fig. 1). This may reflect that 23% of the players on particular Canadian teams in 2011–2012 were members of those teams in 2007–2008. This pattern may also be due to chance. Consistent with both possibilities, the correlation across years for the six common Canadian teams' correlations was positive yet modest ($r(4) = .19$, $p = .72$).

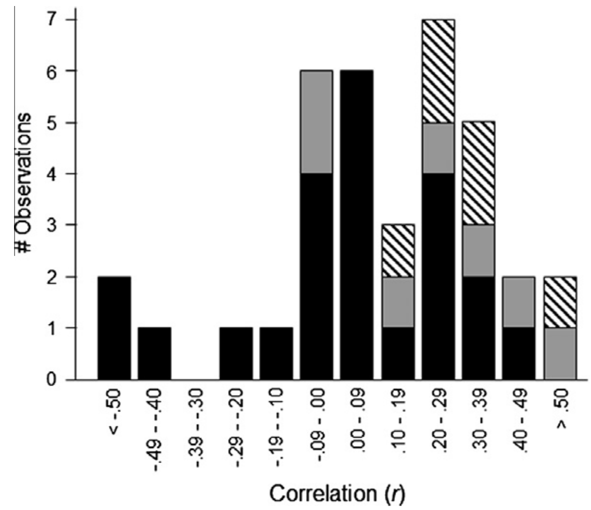

Fig. 1. Histogram showing the frequency of correlation coefficients for the association between FWHR and penalty minutes per game for NHL hockey teams. Black bars indicate correlations for 23 American teams assessed in the present study; gray bars indicate seven Canadian teams in present study; hatched bars indicate six Canadian teams in Carré and McCormick (2008).

Because penalties is skewed heavily to the right (Shapiro–Wilk $W = .81$, $p < .0001$), we repeated our tests after log transforming it. Across all players, FWHR and logpenalties were positively correlated, but this did not reach statistical significance ($r(518) = .076$, $p = .085$). Similarly, the rank-order correlation between FWHR and penalties did not reach significance ($\rho(518) = .078$, $p = .073$).

We next sought to improve the tests by decreasing random error in measures of FWHR and penalties. We did this by excluding players who were smiling, rather than showing neutral facial expressions, and by excluding those who had played fewer than 60 career games. Again, FWHR was not significantly correlated with either penalties ($r(252) = .10$, $p = .11$) or logpenalties ($r(252) = .07$, $p = .30$).

3.1. Fighting

Because overall penalties include fighting penalties, a positive correlation was expected between logpenalties and logfighting, and this was obtained ($r(469) = .67$, $p < .0001$). Nonetheless, if FWHR is truly associated with aggression the correlations should be stronger for FWHR and fighting penalties than for FWHR and overall penalties. Contrary to this, FWHR was not significantly correlated with logfighting ($r(469) = .04$, $p = .41$), and this result did not change when smiling players were excluded ($r(252) = .06$, $p = .34$).

3.2. Body size

FWHR was not correlated with height ($r(518) = -.02$, $p = .64$) or weight ($r(518) = .02$, $p = .73$). Logpenalties was significantly positively correlated with height ($r(518) = .29$, $p < .0001$) and weight ($r(518) = .44$, $p < .0001$). In a multiple regression, these variables collectively predicted substantial variance in logpenalties ($R^2 = .20$, $F(3,516) = 16.2$, $p < .0001$); weight was a significant predictor ($\beta = .48$, $p < .0001$), but height ($\beta = -.06$, $p = .33$) and FWHR ($\beta = .07$, $p = .09$) were not. We repeated this analysis after excluding those who smiled or played fewer than 60 games, and the results were similar (weight: $\beta = .48$, $p < .0001$; height: $\beta = -.02$, $p = .81$; FWHR: $\beta = .07$, $p = .21$).

We addressed the robustness of weight as predictor of penalties in two ways. First, to check if the relation might only be driven by a small group of highly aggressive players (i.e. "enforcers") we removed the two most penalized players on each team; the relation between weight and logpenalties remained substantial ($r(458) = .35, p < .0001$). Second, we computed the correlations separately for each team; 29 of 30 were positive, and 12 of these (all positive) reached significance. Even after removing the two most penalized players, the correlations were positive for 27 of 30 teams and nine of these (all positive) reached significance.

Body size also predicted fighting penalties. Logfighting was significantly positively correlated with height ($r(469) = .18, p < .0001$) and weight ($r(469) = .28, p < .0001$). In a multiple regression with these variables and FWHR, body weight was the only significant predictor ($R^2 = .09$; weight: $\beta = .32$, $p < .0001$; height: $\beta = -.05$, $p = .44$; FWHR: $\beta = .03$, $p = .45$). We repeated this analysis after excluding those who smiled and obtained similar results (weight: $\beta = .38, p < .0001$; height: $\beta = -.10, p = .26$; FWHR: $\beta = .06, p = .34$).

4. Discussion

This study has three principle findings. First, the positive correlation between FWHR and penalties in 520 NHL players was slight ($r = .09$), did not reach statistical significance ($p = .06$), and was significantly smaller than the correlation reported by Carré and McCormick (2008; $n = 112$; $r = .30$). Thus, although FWHR may predict penalties in some teams, it does not do so generally (Fig. 1). We suggest that Carré and McCormick (2008) result represents Type 1 error; in other words, by chance, they happened to select a set of teams that showed the relationship.

Our second main finding was that FWHR did not correlate with another measure of aggression, fighting penalties. This corroborates that FWHR is not associated with aggression in this population.

Our third finding was that body size, particularly body weight, was a robust predictor of aggression, as assessed by both overall penalties and fighting penalties. This finding is consistent previous work associating body size with aggression (e.g., Ishikawa et al., 2001) and strength (e.g., Fink, Neave, & Seydel, 2007; Sell et al., 2009). We anticipated that body size would be a stronger predictor of fighting than overall penalties, but this was not the case. This could reflect that fighting penalties were based on only a single previous season whereas as overall penalties were based on career totals.

Two concerns arise regarding our conclusion that FWHR is unrelated to aggression in this population. First, there was a positive association, and it nearly reached statistical significance ($p = .06$). This raises the possibility of Type 2 error, i.e. studies of other large samples of pro hockey players might reveal a significant effect. This is certainly plausible, but our results indicate that regardless of whether the effect reaches the threshold of significance, the variance explained by FWHR will be very small. In the present study, body weight predicted about 20% of the variation in penalties whereas FWHR predicted less than 1%.

A second concern involves our measurement of FWHR, particularly the fact that the photos, although of high quality, were not produced for the purpose of allowing standardized measurements. Thus, photos varied in facial expression and horizontal and vertical orientation. We addressed this by eliminating obviously turned or angled photos and testing if our results changed when excluding smiling players. Our general results were unaffected, but we acknowledge that it would be desirable to repeat our study using standardized photos. Future studies might also include other measures of facial masculinity and aggressiveness, including subjective judgments.

Besides differing from Carré and McCormick (2008), our results apparently conflict with previous studies linking FWHR with aggression (studies 1 and 2 of Carré & McCormick, 2008; Stirrat & Perrett, 2010 see also Christiansen & Winkler, 1992). Although it is conceivable that all of these findings represent Type 1 error (i.e. false positives), this seems highly unlikely given their convergence. Furthermore, Carré et al. (2010) have provided evidence that facial perceptions of aggression (which predict aggression in laboratory tasks; Carré & McCormick, 2008) are largely based on FWHR. Facial perceptions of aggression, in turn, may be closely associated with facial perceptions of fighting ability and upper body strength, and Sell et al. (2009) have shown that these perceptions closely track measured upper body strength and fighting frequency. Similarly, Fink et al. (2007) have shown that facial perceptions of dominance correlate positively with hand grip strength. Thus, there seems little reason to doubt that FWHR generally predicts aggressiveness and associated attributes (but for another FWHR null finding, see Özener, in press).

Therefore, our unusual finding is best attributed to the fact that our sample—men playing ice hockey in the premier professional league in the world—is extraordinarily selective. In the US, for example, there are approximately 36,000 boys that play in high school (National Federation, 2010) and many others playing for more selective teams, but only 234 American men played at least one NHL game in 2010–2011 (nhl.com), indicating that the odds of "making it" are less than 0.6%. Given that ice hockey requires strength, speed, and the ability to initiate and withstand physical contact, it is likely that almost all NHL players are moderately to highly masculine. Supporting this are FWHR data from Carré and McCormick (2008; Table 1 and Fig. 4); male undergraduates and NHL players differed by more than one standard deviation in FWHR, about twice as large as the sex difference among undergraduates. It seems plausible, then, that all NHL players are somewhat prone to commit aggression, especially because aggression is thought to be instrumental to success in this sport (Bernstein, 2006).

Despite the selectivity of our population, our result should not be interpreted as merely indicating a ceiling effect whereby there is insufficient variation in one or both traits to plausibly allow an association. A ceiling effect can be ruled out because the standard deviations in FWHR for populations of undergraduates and NHL players are similar (Carré & McCormick, 2008), and, in the present study, there was sufficient variation in aggression to allow robust associations with body size.

Facial characteristics are known to influence observers' predictions of behavior in domains ranging from criminal sentencing (Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006) to government elections (Olivola & Todorov, 2010). In such domains, the populations under consideration may be highly selective, and our professional hockey result suggests that the predictions may be completely inaccurate in such contexts, even if they hold across general populations, such as undergraduates. Tendencies to overgeneralize facial cues (Zebrowitz & Montepare, 2008) make sense from an evolutionary perspective because the mechanism(s) underlying such biases may not have been selected against during human evolutionary history when group sizes were smaller than in most modern societies and highly selective groups would have been rare.

We know of two apparent challenges to the hypothesis that inferences based on facial characteristics generally will be inaccurate in highly selective populations. One is demonstrations that facial perceptions of competence and leadership in chief executive officers (CEOs) predict the profitability of the world's leading corporations (e.g., Rule & Ambady, 2009). Nonetheless, there are several interpretations for such patterns, including that companies compete to hire CEOs that "look the part", even if an individual's appearance does not correspond to their ability. Another apparent challenge is a report that professional quarterbacks' passing ratings correlated with their attractiveness (Williams, Park, & Wieling, 2010). However, the effect was small and may not prove robust. In conclusion, the existing evidence appears consistent with the hypothesis that, although observers readily make inferences based on facial appearance, such inferences will be inaccurate when evaluating highly selective populations.

Acknowledgments

We thank Amanda Taylor and Bob Erickson for assistance with measurements and Justin Carré, Mike Lombardo, Cheryl McCormick, David Puts, Stephen Shepherd, Alex Todorov, two anonymous reviewers, and the Editor for comments on the manuscript.

References

Bernstein, R. (2006). *The code: The unwritten rules of fighting and retaliation in the NHL*. Chicago: Triumph Books.
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, B, 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.
Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception, 39*, 356–377.
Christiansen, K., & Winkler, E. M. (1992). Hormonal, anthropometrical, and behavioral correlates of physical aggression in !Kung San men of Namibia. *Aggressive Behavior, 18*, 271–280.
Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy – Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*, 383–386.
Federation, National (2010). *2009–10 High school athletics participation survey*. Indianapolis, IN: National Federation of State High School Associations.
Fink, B., Neave, N., & Seydel, H. (2007). Male facial appearance signals physical strength to women. *American Journal of Human Biology, 19*, 82–87.
Haselhuhn, M. P., & Wong, E. M. (2011). Bad to the bone: Facial structure predicts unethical behavior. *Proceedings of the Royal Society of London, B.*. doi:10.1098/rspb.2011.1193.
Ishikawa, S. S., Raine, A., Lencz, T., Bihrle, S., & LaCasse (2001). Increased height and bulk in antisocial personality disorder and its subtypes. *Psychiatry Research, 105*, 211–219.
Kramer, R. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology, 63*, 2273–2287.
Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior, 34*, 83–110.
Özener, B. (in press). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior*. doi:10.1016/j.evolhumbehav.2011.08.001.
Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles, 61*, 644–652.
Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London, B, 276*, 575–584.
Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*, 349–354.
Weston, E. M., Friday, A. E., & Lio, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *PLoS ONE, 2*, 1–8.
Williams, K. M., Park, J. H., & Wieling, M. B. (2010). The face reveals athletic flair: Better National Football League quarterbacks are better looking. *Personality and Individual Differences, 48*, 112–116.
Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2*, 1497–1517.
Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior, 28*, 167–185.