PSYCHOLOGICAL SCIENCE

A JOURNAL OF THE ASSOCIATION FOR PSYCHOLOGICAL SCIENCE aps

Research Report

Facial Structure Is a Reliable Cue of Aggressive Behavior

Justin M. Carré,[1] Cheryl M. McCormick,[1,2] and Catherine J. Mondloch[1,2]

[1]*Department of Psychology and* [2]*Centre for Neuroscience, Brock University*

ABSTRACT—*Facial width-to-height ratio is a sexually dimorphic metric that is independent of body size and may have been shaped by sexual selection. We recently showed that this metric is correlated with behavioral aggression in men. In Study 1, observers estimated the propensity for aggression of men photographed displaying neutral facial expressions and for whom a behavioral measure of aggression was obtained. The estimates were correlated strongly with the facial width-to-height ratio of the stimulus faces and with the actual aggression of the men. These results were replicated in Study 2, in which the exposure to each stimulus face was shortened to 39 ms. Participants' estimates of aggression for each stimulus face were highly correlated between Study 2 (39-ms exposure) and Study 1 (2,000-ms exposure). These findings suggest that the facial width-to-height ratio may be a cue used to predict propensity for aggression in others.*

The human face is a basis for judgments about gender, ethnicity, attractiveness, emotion, and personality traits (Zebrowitz, 2006). In fact, personality attributions based on the characteristics of the face show high rates of consensus across observers (Penton-Voak, Pound, Little, & Perrett, 2006; Todorov, Mandisodza, Goren, & Hall, 2005), are made very quickly (Bar, Neta, & Linz, 2006; Willis & Todorov, 2006), and, for certain traits, are somewhat accurate (Penton-Voak et al., 2006). People are relatively good at identifying "cheaters" in a Prisoner's Dilemma game based on facial photographs (Verplaetse, Vanneste, & Braeckman, 2007), and women's judgments of men's interest in infants based on their faces predicted their actual interest in infants (Roney, Hanson, Durante, & Maestripieri, 2006). There is some evidence that "baby facedness," characterized by round faces and big eyes, is associated with social, intellectual, and physical weakness (Zebrowitz, Fellous, Mignault, & Andreoletti, 2003). Also, people are accurate in estimating the physical strength and fighting ability of others based on facial information (Sell et al., 2009), although the facial metrics used to make such judgments are not well understood.

Recently, Weston, Friday, and Lio (2007) identified a sexually dimorphic characteristic of the face that was independent of body size from a morphometric analysis of an ontogenetic series of human skulls. They found that the growth trajectories of males and females diverged at puberty for bizygomatic width and not for upper facial height (from the upper lip to the mid-brow), leading to a width-to-height facial dimorphism (greater ratio in men than in women). They proposed that the sexual dimorphism in facial width-to-height may reflect a sexual selection pressure that is independent of selection for body size.

We recently reported that individual differences in the facial width-to-height ratio (WHR) accounted for a significant proportion of variance in aggressive behavior in men, but not women, tested in the laboratory (Carré & McCormick, 2008). Aggression was measured in the lab using a modified version of the point-subtraction aggression paradigm, in which players have continuous access to three buttons: Pressing one button earns points, pressing another protects the player from having points stolen, and pressing the third button steals points from a fictitious opponent to no benefit to the player. Aggression is defined as the number of button presses on that third button. Facial WHR was correlated with actual aggression in men (r = .38), but not women. A similar finding was obtained in elite male hockey players when aggression was defined as the number of penalty minutes per game. The goal of the current study was to determine whether observers' estimates of propensity for aggression are correlated with individual differences in the facial WHR, a finding that would suggest that this facial metric could be used to predict another's propensity for aggressive behavior.

Here, we asked observers to estimate the propensity for aggression of men photographed displaying neutral facial expressions and for whom a behavioral measure of aggression previously had been obtained (Carré & McCormick, 2008). In the first study, participants viewed the stimulus faces for 2,000 ms; in the second study, participants viewed the stimulus faces for 39 ms.

Address correspondence to Cheryl M. McCormick, Department of Psychology and Centre for Neuroscience, Brock University, 500 Glenridge Ave., St. Catharines, Ontario, Canada L2S 3A1, e-mail: cmccormick@brocku.ca.

RESEARCH, THEORY, & APPLICATION IN PSYCHOLOGY AND RELATED SCIENCES
Psychological Science

METHOD

Participants

Two samples of undergraduate students (Study 1: 16 women, 15 men; mean age = 19.94 years, SD = 2.05 years; Study 2: 16 women, 0 men; mean age = 19.38 years, SD = 1.41 years) received course credit for participation. The procedures were approved by Brock University's Research Ethics Board.

Stimuli

Photographs were obtained from a sample of 37 men for which aggressive behavior and facial WHR was quantified previously (see Carré & McCormick, 2008). These men were volunteers from an introductory psychology participant pool who received a $5 honorarium and course credit for their participation. Aggressive behavior was measured using a modified version of the point subtraction aggression paradigm, a well-validated laboratory task (e.g., Cherek, Schnapp, Moeller, & Dougherty, 1996). The facial WHR of the men was measured using NIH ImageJ software and involved the landmarks originally used by Weston et al. (2007). Specifically, the distance between the left and right zygion (bizygomatic width) was divided by the distance between the upper lip and mid-brow (upper facial height) to yield the facial WHR (see Fig. 1). For the studies described here, the sample of stimulus faces was reduced to include only Caucasian men without facial hair (to avoid judgments based on stereotypes) and displaying neutral expressions (n = 24, mean age = 19.08 years, SD = 1.41 years). Faces were converted to 8-bit gray scale, standardized using a hairline-chin distance of 400 pixels, and placed within a black background.

Procedure

In the first study, stimulus faces were presented on a black background using E-Prime software on a 14-in. LCD monitor.

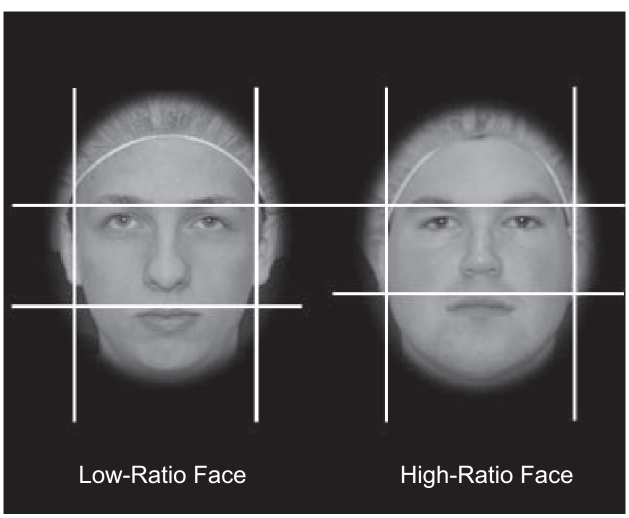

Fig. 1. Examples of stimuli used in Studies 1 and 2. The faces differ in width-to-height ratio (i.e., high and low ratios). The lines drawn on the faces were not shown to observers, and are included here to illustrate the landmarks used to measure the width-to-height ratio.

Images were approximately 17 cm wide by 20 cm high (or 15.2 × 12.9 visual degrees when viewed from 75 cm). Faces were presented in random order for 1,000 ms to familiarize the participants with the range of faces. Participants were told how the aggressive behavior of the men had been assessed. Next, each face was presented for 2,000 ms (fully randomized), after which the question "How aggressive would this person be if provoked?" appeared on a black background along with a 7-point Likert scale (1 = *not at all aggressive*, 7 = *very aggressive*). Participants were given unlimited time to make their response on a numerical keypad, which then caused the next stimulus face to appear. After completing the estimates of aggression, participants rated each face for dominance, masculinity, trustworthiness, and attractiveness using 7-point Likert scales (order of these four ratings was fully counterbalanced). In sum, each face was presented a total of six times, once for familiarization and once for each of the five traits, and the entire set of faces was rated on one attribute at a time. Estimates of aggression were highly correlated with ratings of dominance (r = .92), masculinity (r = .86), trustworthiness (r = −.90), and attractiveness (r = −.57). Only estimates of aggression and of dominance were correlated significantly with the facial WHR (r = .59 and r = .54, respectively). However, given the specific evaluative context (i.e., participants were told of the objective measure of aggression; Oosterhoff & Todorov, 2008) and the high degree of association of the five ratings, statistical analyses used estimates of aggression only.

In the second study, presentation of a stimulus face was preceded by a central fixation cross that appeared for 500 ms. A face was then presented for 39 ms (order of faces was fully randomized and without any prior familiarization to the faces), after which the question "How aggressive would this person be if provoked?" appeared on a black background along with a 7-point Likert scale (1 = *not at all aggressive*, 7 = *very aggressive*). As in Study 1, participants were told how the aggressive behavior of the men had been assessed before the presentation of stimulus faces.

Statistics

Cronbach's alpha was calculated to examine the consistency of the ratings of estimated aggression across individual participants. For each participant within each study, we calculated the correlation between the estimate of aggression for the 24 faces and both the facial WHR and actual aggression of the stimulus faces. One-sample t tests were computed to test the primary hypothesis that these correlations would be significantly different from the null hypothesis (i.e., no association). For each stimulus face, we calculated the mean estimated aggression across participants and correlated that with both the facial WHR and the actual aggression of the stimulus face. The correlation between estimates of aggression in Study 1 and Study 2 was also calculated for each face. Significance level was set at $p < .05$, two-tailed, for all analyses.

RESULTS

Study 1

The estimates of aggression were highly consistent across individual observers (Cronbach's $\alpha = .95$). Single-sample t tests comparing individual correlations to a null value of zero showed that estimated aggression was positively associated with the facial WHR of the stimulus faces: male observers, $t(14) = 16.94$, $p < .001$, $p_{rep} > .99$; female observers, $t(15) = 9.23$, $p < .001$, $p_{rep} > .99$; combined, $t(30) = 16.41$, $p < .001$, $p_{rep} > .99$ (Fig. 2). These t tests also indicated that estimated aggression was positively associated with actual aggression of the stimulus faces: male observers, $t(14) = 6.95$, $p < .001$, $p_{rep} > .99$; female observers, $t(15) = 8.81$, $p < .001$, $p_{rep} > .99$; combined, $t(30) = 11.21$, $p < .001$, $p_{rep} > .99$ (Fig. 2). The mean estimated aggression for each face across participants was associated with both the facial WHR ($r = .59$, $p < .002$, $p_{rep} = .98$) and actual aggression ($r = .42$, $p = .04$, $p_{rep} = .89$) of the stimulus faces (Fig. 2).

Study 2

The estimates of aggression were highly consistent across individual observers (Cronbach's $\alpha = .89$). Also, estimates of aggression among these participants were highly correlated with estimates of aggression from participants in Study 1, who were given 2,000-ms exposure to the stimulus faces, $r = .82$, $p < .001$, $p_{rep} > .99$ (Fig. 3). Single-sample t tests comparing individual correlations to a null value of zero showed that estimated aggression was positively associated with both the facial WHR, $t(15) = 10.24$, $p < .001$, $p_{rep} > .99$, and actual aggression of the stimulus faces, $t(15) = 4.49$, $p < .001$, $p_{rep} > .99$. The mean estimated aggression for each face across participants was associated with the facial WHR but not with actual aggression, $r = .70$, $p < .001$, $p_{rep} > .99$, and $r = .31$, $p = .14$, $p_{rep} = .79$ (Fig. 4).

DISCUSSION

The results indicate that observers can make accurate judgments of propensity for aggression from faces displaying neutral

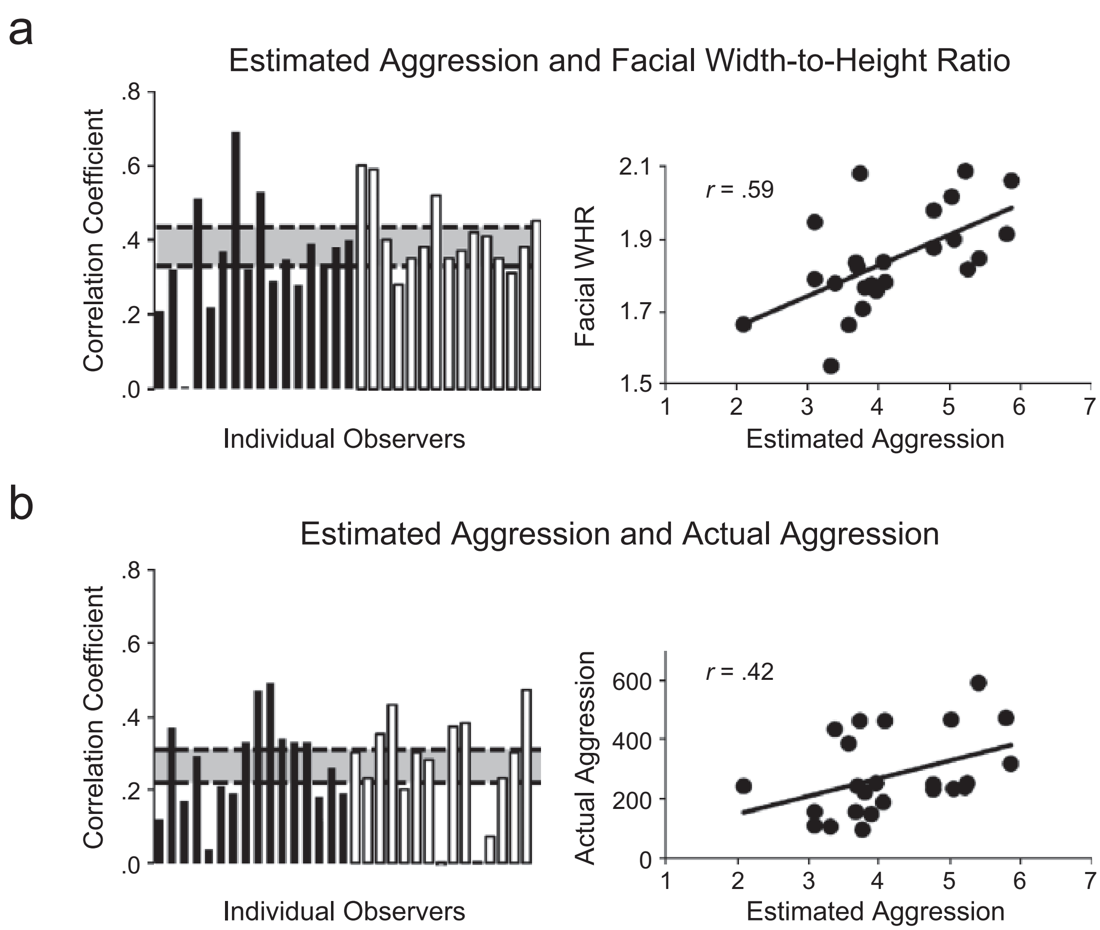

Fig. 2. Results from Study 1: relationships between observers' estimates of the pictured individuals' aggression and (a) width-to-height ratio (WHR) of the facial stimuli and (b) the pictured individuals' actual aggression. The bar graphs show the relationships for each individual observer (stimuli viewed for 2,000 ms each). Black bars indicate female observers ($n = 16$); white bars indicate male observers ($n = 15$). Shaded areas represent the 95% confidence intervals (CIs). The mean r value for the relationship between estimated aggression and WHR was .38 (95% CI = .33–.43), and the mean r value for the relationship between estimated aggression and actual aggression was .27 (95% CI = .21–.31). The graphs on the right show facial WHR and actual aggression as a function of estimated aggression, across participants.

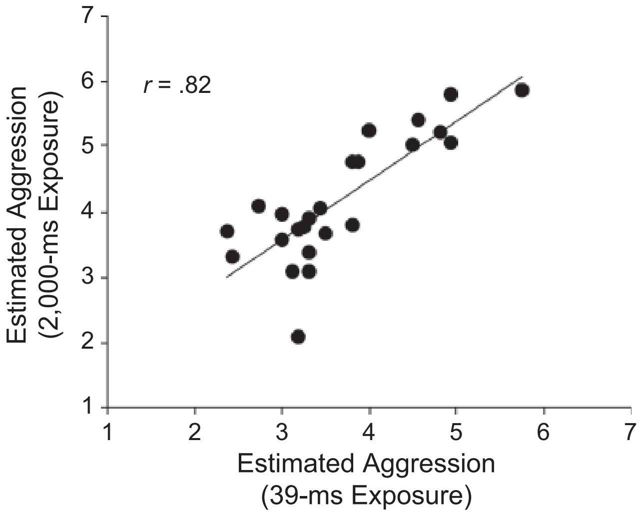

Fig. 3. Relationship between observers' estimates of the pictured individuals' aggression in Study 1 (2,000-ms exposure) and in Study 2 (39-ms exposure). The plotted points represent individual stimulus faces; each point shows the average estimate of aggression for a stimulus across participants in Study 1 and across participants in Study 2.

expressions, even when exposure to the faces is limited to 39 ms. In both Study 1 and Study 2, individual participants reliably judged men with larger facial WHRs as more aggressive; across participants, faces with larger WHRs were rated as more aggressive than faces with smaller WHRs. The strong correlation between estimated aggression and facial WHR suggests that this facial metric may be one of the facial cues used to make accurate estimates of aggression.

It is not surprising that participants' estimates of aggression were correlated more strongly with facial WHR than with actual aggression. Whereas facial WHR is a stable facial characteristic that provides a static estimate of the propensity for aggression, any actual behavioral aggression will vary over time (e.g., as a function of state or situation). Indeed, it is impressive that the correlation between estimates of aggression and actual aggression of the stimulus faces was as high in the current study (r = .42 and r = .31) as was the correlation between facial WHR and actual aggression in our original study (r = .38; Carré & McCormick, 2008). Furthermore, facial WHR is only one of many cues to propensity for aggression.

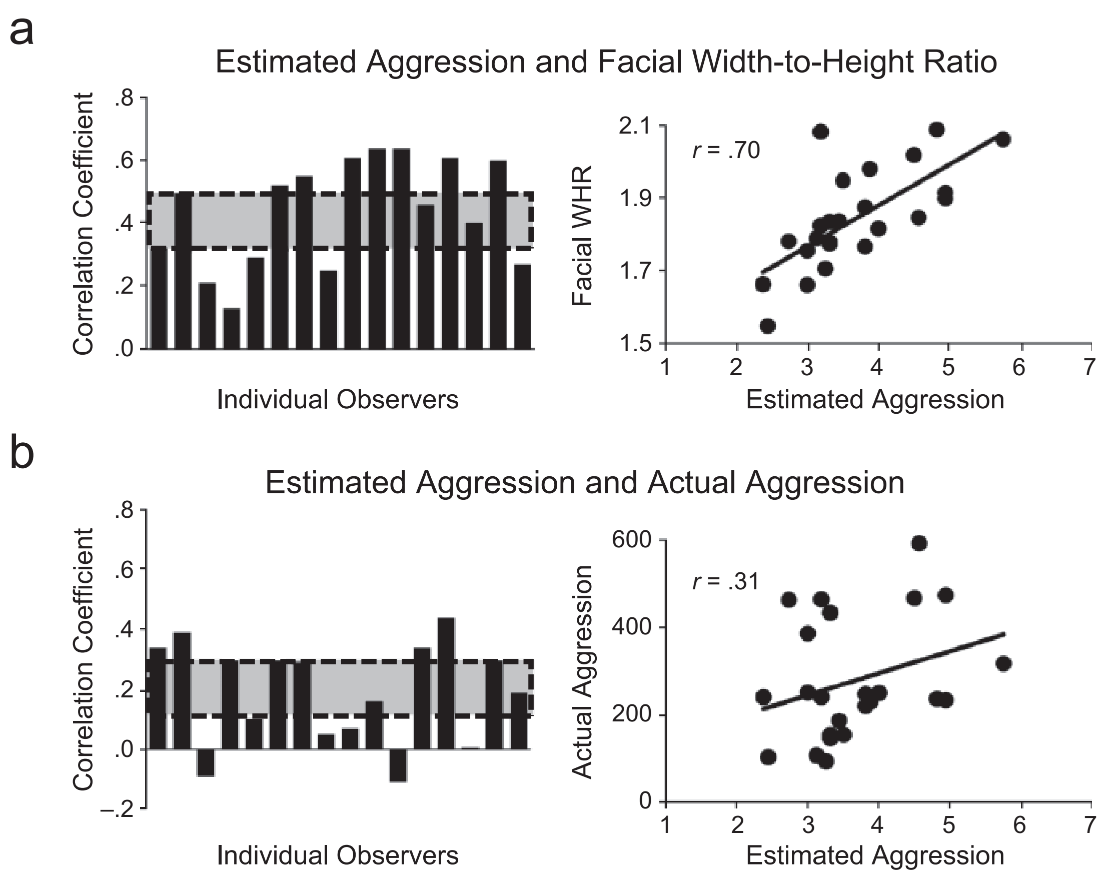

Fig. 4. Results for Study 2: relationships between observers' estimates of the pictured individuals' aggression and (a) facial width-to-height ratio (WHR) and (b) the pictured individuals' actual aggression. The bar graphs show the relationship for each individual observer (n = 16; stimuli viewed for 39 ms). Shaded areas represent the 95% confidence intervals (CIs). The mean r value for the relationship between estimated aggression and WHR was .44 (95% CI = .35–.53), and the mean r value for the relationship between estimated aggression and actual aggression was .19 (95% CI = .10–.28). The graphs on the right show facial WHR and actual aggression as a function of estimated aggression, across participants.

Future research should investigate whether the facial WHR meets the criteria for an "honest signal" of aggressive potential, similar to honest signals guiding interindividual behavior in other species (Setchell, Smith, Wickings, & Knapp, 2008; Tibbetts & Dale, 2004). Honest signals are used in other species as a means to gauge one's relative status within the hierarchy (Setchell et al., 2008; Tibbetts & Dale, 2004) and may serve to modulate adaptive behavior (Senar & Camerino, 1998; Tibbetts & Lindsay, 2008). The ability to gauge aggressive behavior from neutral faces may reflect an overgeneralization of emotional expressions (Montepare & Dobish, 2003). In other words, neutral faces may be evaluated according to their similarity to certain emotional expressions, such as anger and happiness, which may in turn be used by perceivers to guide adaptive social behavior (e.g., approach/avoidance; Oosterhof & Todorov, 2008). Notably, angry facial expressions consist of lowering the brow and raising the upper lip, a facial movement that inevitably increases the facial WHR and, by implication, increases the saliency of the "signal" advertising propensity for aggression. Thus, it is also possible that the relationship between facial WHR and aggression reflects social conditioning whereby a person's aggressive behavior has been shaped by others' expectations of their aggressive behavior. Furthermore, it may be some other cue in the face correlated with the facial WHR that is influencing estimates of aggression. For example, using computer-generated faces, Todorov, Baron, and Oosterhof (2008) found that variation in the brow ridge, cheekbones, chin, and nose sellion (i.e., where the nose and brow meet) were related to observer ratings of trustworthiness, a social attribute that is negatively correlated with aggression. Thus, it will be important to test whether the facial WHR predicts aggression independently of these other cues. Nevertheless, the present results raise the possibility that subtle differences in facial structure influence trait judgments, which may, in turn, guide social behavior.

Acknowledgments—We thank Mark Vida and Alexandra Hatry for technical assistance. C.M.M. and C.J.M. hold Natural Sciences and Engineering Research Council (NSERC) Discovery Grants. J.M.C. holds an NSERC Canada Graduate Scholarship.

REFERENCES

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 268–278.

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: Biological Sciences*, *275*, 2651–2656.

Cherek, D., Schnapp, W., Moeller, F., & Dougherty, D. (1996). Laboratory measures of aggressive responding in male parolees with violent and non-violent histories. *Aggressive Behavior*, *22*, 27–36.

Montepare, J.M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*, 237–254.

Oosterhof, N.N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, USA*, *105*, 11087–11092.

Penton-Voak, I.S., Pound, N., Little, A.C., & Perrett, D.I. (2006). Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception. *Social Cognition*, *24*, 490–524.

Roney, J.R., Hanson, K.N., Durante, K.M., & Maestripieri, D. (2006). Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society B: Biological Sciences*, *273*, 2169–2175.

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 B: Biological Sciences*, *276*, 575–584.

Senar, J.C., & Camerino, M. (1998). Status signalling and the ability to recognize dominants: An experiment with siskins (*Carduelis spinus*). *Proceedings of the Royal Society B: Biological Sciences*, *265*, 1515–1520.

Setchell, J.M., Smith, T., Wickings, E.J., & Knapp, L.A. (2008). Social correlates of testosterone and ornamentation in male mandrills. *Hormones & Behavior*, *54*, 365–372.

Tibbetts, E.A., & Dale, J. (2004). A socially enforced signal of quality in paper wasps. *Nature*, *432*, 218–222.

Tibbetts, E.A., & Lindsay, R. (2008). Visual signals of status and rival assessment in *Polistes dominulus* paper wasps. *Biology Letters*, *4*, 237–239.

Todorov, A., Baron, S.G., & Oosterhof, N.N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience*, *3*, 119–127.

Todorov, A., Mandisodza, A.N., Goren, A., & Hall, C.C. (2005). Inferences of competence from faces predict election outcome. *Science*, *308*, 1623–1626.

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. A kernel of truth in predicting cheating detection. *Evolution and Human Behavior*, *28*, 260–271.

Weston, E.M., Friday, A.E., & Lio, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *PLoS ONE*, *2*, 1–8.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science*, *17*, 592–598.

Zebrowitz, L.A. (2006). Finally, faces find favour. *Social Cognition*, *24*, 657–701.

Zebrowitz, L.A., Fellous, J.M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*, 194–215.

(Received 1/19/09; Revision accepted 3/19/09)