The face reveals athletic flair: Better National Football League quarterbacks are better looking

Kevin M. Williams\textsuperscript{a,}\textsuperscript{*}, Justin H. Park\textsuperscript{b}, Martijn B. Wieling\textsuperscript{c}

\textsuperscript{a}Multi-Health Systems, Inc., Toronto, Canada M2H 3M6
\textsuperscript{b}Department of Experimental Psychology, University of Bristol, Bristol, United Kingdom
\textsuperscript{c}Department of Computational Linguistics, University of Groningen, Groningen, The Netherlands

\section*{A B S T R A C T}

We investigated whether individual differences in athleticism, and in turn heritable fitness, may be signaled by differences in facial attractiveness within a highly select group of athletes: National Football League (NFL) quarterbacks (QBs). Athleticism was operationalized as the passer rating, the NFL's official measure of performance among QBs. Results from a preliminary study showed a positive correlation between 30 NFL QBs' passer ratings and their facial attractiveness as rated by 30 women. In a further study, a different group of 30 women rated a different cohort of 58 NFL QBs. The results showed that the QBs' mean attractiveness ratings were positively correlated with their passer ratings, which was found to be independent of players' age, ethnicity, height, weight, or facial expression. These findings build upon previous research and provide further support for the hypothesis that individual differences in athleticism, a heritable trait desirable in mate selection, may be signaled reliably through facial attractiveness.

\section*{1. Introduction}

Identifying indicators of heritable fitness is valuable, as they inform our understanding of the human mate-selection process and other affiliative behaviors. Heritable fitness is associated with immunocompetence and the ability to cope with stressors; consequently, animals that seek out mates with high heritable fitness increase the likelihood of obtaining 'good genes' for their offspring (Gangestad & Simpson, 2000; Geary, 1998). Individuals high in heritable fitness advertise their underlying genetic quality via phenotypic fitness indicators, and potential mates are attracted to these indicators. For example, peaks are attracted to peacocks with larger and more colorful trains because maintenance of trains requires large amounts of caloric energy and resistance to environmental stressors. In turn, it is beneficial for the female to pass these traits onto their offspring.

Here, we focused on athleticism as a trait indicating heritable fitness in humans. Not only is athleticism sexually attractive, the tangible benefits of athleticism are apparent. Throughout human evolutionary history, athleticism and physical superiority would have been vital in situations such as hunting and physical competition for mates and other valued commodities, particularly among males. It has moreover been argued that sports – a modern-day expression of physical competition – are systems designed to advertise these desirable physical traits (Miller, 2000). Consistent with this idea is evidence that various athletic traits are highly heritable (e.g., Missitzi, Geladas, & Klissouras, 2004).

An important physiological mediator of athleticism is testosterone. Indeed, testosterone is widely considered a key contributor to athletic performance in both men and women through its effects on the brain and/or vascular system (see Hönekopp, Manning, & Müller, 2006), and it is highly heritable (Harris, Vernon, & Boomsma, 1998). In turn, testosterone-linked traits are often hypothesized to be fitness indicators (Hoekstra, Bartels, & Boomsma, 2006). For instance, there is evidence that masculinity – which is associated with testosterone – serves as a fitness indicator (Frederick & Haselton, 2007).

Given the influence of testosterone on general physical characteristics, it is plausible that differences in testosterone are also reflected in facial features. For example, ‘masculinized’ faces (i.e., larger jaw, more prominent brow ridge) may be related to higher testosterone levels. Penton-Voak and Perrett (2000) found that masculinized faces are rated as more attractive by females when in the most fertile stages of the menstrual cycle (cf. Perrett et al., 1998; Rhodes, Hickford, & Jefferies, 2000). Other research suggests that facial attractiveness more generally serves as a fitness indicator (Perrett et al., 1999; Rhodes, Proffitt, Grady, & Sumich, 1998; Scheib, Gangestad, & Thornhill, 1999).
Given that both athleticism and facial attractiveness appear to be linked to heritable fitness, a plausible hypothesis is that there is a correlation between athleticism and facial attractiveness. This hypothesis was the focus of the present research.

Facial attractiveness ratings are often used in attractiveness studies, given that they can be rated quickly (e.g., Olsen & Marshuetz, 2005). Developmentally, evaluations of facial attractiveness appear to be among the most elementary human processes, as even infants show an ability to make such differentiations (Langlois, Ritter, Roggmann, & Vaughn, 1991). Also, previous research has established that individuals are able to evaluate physical strength through facial cues (Sell et al., 2009).

Are more athletic individuals in fact perceived to be more attractive when only their faces are rated? Some intriguing results were reported by Park, Buunk, and Wieling (2007). These authors reasoned that within some team sports, certain positions (e.g., soccer goalkeepers) may require higher degrees of athleticism (a set of traits closely associated with heritable fitness). According to this line of reasoning, players at these positions should be rated as more facially attractive than players at other positions. Park and colleagues examined women's facial attractiveness ratings of professional male European football (soccer) and ice hockey players. Within each sport, players at positions hypothesized to require more athleticism were rated as most attractive. Although intriguing, this finding is limited by the fact that players across different positions were compared; the interpretation hinges on the assumption that certain positions actually require greater athleticism. To investigate whether more athletically gifted players have more attractive faces, a more objective criterion for athleticism is necessary. Thus, in the present research, we focused on athletic performance within a single position.

In the current context, the National Football League (NFL) provides a fitting sample of research subjects. It has been argued that the more prestigious sports are those that more honestly signal traits (e.g., strength, endurance, agility, and intelligence) that are associated with heritable fitness (Miller, 2000). The NFL is widely considered the most prestigious sport in North America, as evidenced by game attendance, television ratings, and the value of broadcasting rights. Based on current team salary cap figures, which are a proportion of the league's total revenue, it is estimated that the NFL generated over 6.2 billion USD in total revenue in 2008.

In the present research, we focused on the quarterback (QB) position, relying on an objective measure of athletic performance. QB performance was hypothesized to be associated with facial attractiveness. This study thus provided a more rigorous test of the athleticism–attractiveness link.

QB is widely considered the most important position in North American football. The QB is responsible for determining the offensive play, communicating that play to the other offensive players, identifying the opposing team’s defensive strategy, and executing the play. In the NFL, the primary measure of a QB’s overall performance is the passer (or QB) rating. This statistic combines various indicators of the QB’s passing ability (e.g., completion percentage, touchdowns per pass attempt; for a complete description see White & Berry, 2002) into a standardized measure of overall performance. The QB rating has been described as “the official measurement of a quarterback’s performance” (Byrd & Ustler, 2007; p. 8; see also White & Berry, 2002) and “by far the most widespread measure used to rank and differentiate quarterbacks” (White & Berry, 2002; p. 10). Because this statistic does not include accomplishments such as games won or championships, which are determined not solely by the QB but the entire team and other factors (e.g., Berri, Schmidt, & Brook, 2006a), it may be considered a true measure of individual performance. Several researchers have used the QB rating to examine the relative performance of NFL QBs (e.g., Leeds & Kowalewski, 2001; Murrell & Curtis, 1994; Niven, 2005). In sum, the QB rating provides a well-recognized and common metric with which the QBs in our study may be compared. We hypothesized that the passer rating would be associated with attractiveness ratings.

We first present results from a preliminary study suggesting a link between QB rating and facial attractiveness. We then present results from another study, which examined the correlation between QB rating and facial attractiveness while controlling for potential third variables.

2. Preliminary study

2.1. Method

Photos of 30 QBs who played at least four games in the 1997 NFL season served as stimuli for the preliminary study. Photos were collected through various internet search engines. Thirty female students at the University of Groningen voluntarily participated by responding to an e-mail message requesting them to rate the attractiveness of male faces. On their personal computers, participants rated each of the 30 faces on a 10-point scale (1 = very unattractive, 10 = very attractive). For each player, the mean of the 30 ratings was used as their overall attractiveness rating.

The measure of athletic performance was each player’s career QB rating. Possible values for this rating range from 0 to 158.3, with higher values indicating greater performance. These statistics were collected from the official NFL website.

3. Results and discussion

A correlation between QB ratings and mean attractiveness ratings was conducted to determine whether athletic performance could in fact be assessed by simple examination of the QBs’ faces. A one-tailed test of significance was used to reflect the hypothesized link between attractiveness and performance, based on results of previous research (Park et al., 2007). Results demonstrated that attractiveness and QB ratings were positively correlated, r = .31, p < .05, exhibiting a small-to-medium effect size. This result suggested that the association requires further exploration, including its generalization to other cohorts of players and the possible effects of various extraneous variables.

4. Full study

In this study, a different sample of 58 NFL QBs was used, and a different group of 30 raters recruited.

4.1. Method

4.1.1. Stimuli

All 58 QBs whose photos served as stimuli were on active NFL rosters as of the beginning of the 2007 NFL season. Photos were obtained randomly from reputable sports-related websites. All photos were roughly 120 x 180 pixels in size and depicted only the players’ faces.

4.1.2. Participants and procedure

Thirty female students at the University of Groningen voluntarily participated by responding to an e-mail message requesting them to rate the attractiveness of male faces. On their personal computers, participants rated each of the 58 faces on a 10-point scale (1 = very unattractive, 10 = very attractive). For each player, the mean of the 30 ratings was used as their overall attractiveness rating.

4.1.3. Measures of athletic performance and demographic variables

The measure of athletic performance was each player’s career QB rating. Several demographic variables – height, weight, age, and ethnicity (coded through visual inspection of skin color as black = 0, white = 1) – were collected for each player. It was necessary to control for the effects of these variables given that several of them are related to ratings of attractiveness (e.g., Nettle, 2002; Swami, 2006; Tatarunaite, Playle, Hood, Shaw, & Richmond, 2005; Weeden & Sabini, 2005).

To ensure a sufficiently large sample of photos, it was necessary to include photos in which there was some variation in facial expression (i.e., smiling vs. not smiling). Because smiling faces are perceived as more attractive than non-smiling faces (Otta, Abrosio, & Hoshino, 1996), we generated a variable rating the degree to which the QBs were smiling in their photo. Prior to commencement of the analyses, five raters – blind to the hypotheses of the study – independently categorized each of the faces as smiling (1) or not smiling (0). The sum of these ratings was then used as the ‘smiling’ score for each photo. Therefore, photos for which the raters completely agreed as to whether the player was smiling or not obtained the highest (5) or lowest (0) possible score on this variable, respectively. The five raters agreed completely on facial expression for 47 (81.0%) of the photos, and the intraclass correlation (ICC2; average ratings) of the smiling ratings was .94.

5. Results

5.1. Demographic variables and passer ratings

Correlations and descriptive statistics for all variables are listed in Table 1. The career passer ratings of the 58 QBs ranged from 51.8 to 139.4, with a mean of 81.0, very similar to the mean career passer rating of all quarterbacks from the 2007 NFL season (83.5). Roughly 83% of the QBs were of European (white) ethnicity, with the remaining 17% African–American (black).

5.2. Attractiveness ratings

The intraclass correlation coefficient (ICC2) for attractiveness ratings was .95, suggesting high agreement across raters in their evaluations of the photos. Combined across the 30 raters, the attractiveness ratings of the 58 QBs ranged from 2.90 to 7.23 with a mean of 4.81 (SD = .99).

5.3. Passer rating, physical attributes, and attractiveness

Attractiveness ratings were analyzed via hierarchical multiple regression. Passer rating was the primary independent variable of interest, with the various demographic variables included as covariates. One-tailed significance tests are used to reflect our directional hypotheses and previous research findings (e.g., Otta et al., 1996).

Results of the first step of regression analyses are reported in Table 2. In the first step of the hierarchical regression, the demographic variables were entered simultaneously with attractiveness ratings as the outcome. Together, these variables showed a small to medium effect size in predicting attractiveness ($R^2 = .11, p < .05$; $f^2 = .12$). Individually, the variables age ($b = -.31, p < .05$, one-tailed) and smiling ($b = .25, p < .05$, one-tailed) showed independent associations with attractiveness ratings, demonstrating small-to-medium effect sizes.1 Passer rating was added in the second step of the regression analysis to evaluate its contribution to attractiveness ratings independent of the demographic variables (Table 3). After accounting for the variables in step 1, passer rating resulted in a significant increase in the overall model ($\Delta R^2 = .05, p < .05$, one-tailed), reflecting a small-to-medium effect size ($f^2 = .06$). Specifically, passer rating was positively associated with attractiveness ratings, $b = .23, p < .05$, one-tailed, demonstrating a small-to-medium effect size.

6. General discussion

This research showed that more athletic QBs have more attractive faces. This finding supports the hypothesis that facial attractiveness signals heritable fitness, and it adds to research showing that variation in heritable fitness can be detected even among professional athletes (Park et al., 2007). Importantly, the present results provide a clearer demonstration of the athleticism–attractiveness link, as the measure of athleticism was an objective assessment of players within a single position, rather than different positions per se. Moreover, we replicated this finding across two studies.

These results are consistent with the framework described above in which testosterone is seen to play a key role. Although our research does not allow us to draw any conclusions regarding testosterone, it is possible that QBs with higher passer ratings have higher testosterone levels, which may be associated with facial features that women find attractive. The role of testosterone, as well as other mediators, is something that requires further research attention.

Our research does include some limitations. One involves our measure of athletic performance among QBs. Despite its status as a “gold standard” measure, the passer rating is not without its...
Forming QBs? This is a question for future research. More than mediate the differences found in our results remain unidenti-
the size of any correlations with athleticism.

Our findings present several avenues for future research. The factors such as ‘athletic intelligence’ or decision-making skills may improve the veracity of our results. QBs may improve the veracity of our results.

The face reveals athletic flair: Better National Football League quarterbacks are better looking. Critics. Some scholars have argued that the passer rating is unintuitive, unscientific, overly complicated, and ignores certain aspects of player performance such as running ability (Berri, Schmidt, & Brook, 2006b). Berri, Schmidt, and Brook (2006a) take the argument even further, contending that, unlike most other sports, the success of each NFL player is influenced so much by their teammates that a valid measure of any given player’s unique value is virtually impossible.

Recently developed alternative measures of QB performance (Berri et al., 2006b; Joyner et al., 1998; White & Berry, 2002) are fairly complex and often overlap highly with traditional measures. Despite the attractive features of these measures, they have yet to reach the official status of the passer rating and require further research before their unique and incremental value can be determined conclusively. Other alternatives may be too heavily influenced by team performance (e.g., games won, championships) or subjectively awarded (e.g., Pro Bowl nominations). Measures of specific athletic traits such as speed, ability, and strength would be ideal, but these are difficult to obtain. For example, although the annual NFL Scouting Combine measures these traits in prospective NFL players through exercises such as the 40-yard dash, 225-lb bench press, and vertical jump, most quarterbacks do not partici-
pate in all of these events (at the 2009 event, only 4 of 21 QBs partici-
pated in the bench press, for instance). Rather than athleticism, factors such as ‘athletic intelligence’ or decision-making skills may also merit discussion. Finally, a larger sample of both raters and QBs may improve the veracity of our results.

Our findings present several avenues for future research. The athleticism–facial attractiveness link has yet to be studied in athletes. Alternately, the relevance of symmetrical facial features (Gangestad & Simpson, 2000) or differences in facial height-to-width ratio (Carré & McCormick, 2008) could also be examined. The characteristics of the female rats could be examined as well. Given that women seem to be more sensitive to fitness indicators while in the fertile phase of the menstrual cycle, the correlation between athleticism and facial attractiveness may be especially high when the female rates are in the fertile phase.

More broadly, the present research demonstrates the value of an evolutionary perspective in generating new findings. Building upon theory and past research, we predicted – and found – that more athletic QBs have more attractive faces.

Table 2
Full Study hierarchical regression results – Step 1.

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smiling</td>
<td>.11</td>
<td>.07</td>
<td>.22</td>
<td>1.59</td>
<td>.21</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>-.04</td>
<td>.41</td>
<td>-.02</td>
<td>.10</td>
<td>.01</td>
</tr>
<tr>
<td>Age (years)</td>
<td>-.07</td>
<td>.04</td>
<td>-.28</td>
<td>1.95</td>
<td>.25</td>
</tr>
<tr>
<td>Height (inches)</td>
<td>.02</td>
<td>.12</td>
<td>.03</td>
<td>.19</td>
<td>.03</td>
</tr>
<tr>
<td>Weight (pounds)</td>
<td>.00</td>
<td>.01</td>
<td>-.04</td>
<td>.24</td>
<td>.03</td>
</tr>
</tbody>
</table>

Note: N = 58 quarterbacks. Dependent variable = attractiveness ratings. Ethnicity coded as 0 = black, 1 = white. Values at least .10 represent small effect sizes. Values in bold significant at p < .05, one-tailed.

Table 3
Full Study hierarchical regression results – Step 2.

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passer rating</td>
<td>.02</td>
<td>.01</td>
<td>-.23</td>
<td>1.73</td>
<td>.23</td>
</tr>
<tr>
<td>Smiling</td>
<td>.13</td>
<td>.07</td>
<td>.25</td>
<td>1.81</td>
<td>.23</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>-.05</td>
<td>.40</td>
<td>-.02</td>
<td>.13</td>
<td>.02</td>
</tr>
<tr>
<td>Age (years)</td>
<td>-.08</td>
<td>.03</td>
<td>-.31</td>
<td>2.21</td>
<td>.28</td>
</tr>
<tr>
<td>Height (inches)</td>
<td>.02</td>
<td>.12</td>
<td>.03</td>
<td>.18</td>
<td>.02</td>
</tr>
<tr>
<td>Weight (pounds)</td>
<td>.01</td>
<td>.01</td>
<td>-.08</td>
<td>.42</td>
<td>.06</td>
</tr>
</tbody>
</table>

Note: N = 58 quarterbacks. Dependent variable = attractiveness ratings. Ethnicity coded as 0 = black, 1 = white. Values at least .10 represent small effect sizes. Values in bold significant at p < .05, one-tailed.

References


