Attractiveness is positively related to World Cup performance in male, but not female, biathletes

Tim W. Fawcett[1,2]*, Jack Ewans[2], Alice Lawrence[2] and Andrew N. Radford[2]

[1]*Centre for Research in Animal Behaviour (CRAB), University of Exeter*

[2]*School of Biological Sciences, University of Bristol*

*Correspondence: Tim Fawcett, Centre for Research in Animal Behaviour (CRAB), Washington Singer Laboratories, University of Exeter, Exeter EX4 4QG. Phone: +44 1392 725273. E-mail: t.w.fawcett@exeter.ac.uk.

SHORT TITLE: Successful male biathletes are more attractive

Ethics. The procedure was reviewed and approved by the University of Bristol Research Ethics Committee (ref. 12741) and informed consent was given by all participating raters.

Funding. No direct funding for this research.

Competing interests. We have no competing interests.

Authors' contributions. TWF, JE, AL and ANR designed the research; JE and AL collected the data; TWF analysed the data and wrote the paper with input from all other authors.

Acknowledgements. We thank Ian Penton-Voak for discussion and help in setting up E-Prime, and Louise Barrett, Erik Postma, Robbie Wilson and an anonymous reviewer for feedback on an earlier draft.

Data accessibility. Analyses reported in this article can be reproduced using the data and R code provided by Fawcett et al. (in press).

Lay summary

Performance in winter sports predicts attractiveness in men, but not in women. We examined the relationship between career-best performance metrics and attractiveness ratings for men and women who compete annually in the biathlon World Cup, a multidisciplinary sport that combines target shooting and cross-country skiing. Male biathletes who had achieved a higher peak performance in their career were rated as more attractive by the opposite sex, whereas there was no such relationship for female biathletes.

Attractiveness is positively related to World Cup performance in male, but not female, biathletes

Whole-organism performance capacity is thought to play a key role in sexual selection, through its impacts on both intrasexual competition and intersexual mate choice. Based on data from elite sports, several studies have reported a positive association between facial attractiveness and athletic performance in humans, leading to claims that facial correlates of sporting prowess in men reveal heritable or non-heritable mate quality. However, for most of the sports studied (soccer, ice hockey, American football and cycling) it is not possible to separate individual performance from team performance. Here, using photographs of athletes who compete annually in a multi-event World Cup, we examine the relationship between facial attractiveness and individual career-best performance metrics in the biathlon, a multidisciplinary sport that combines target shooting and cross-country skiing. Unlike all previous studies, which considered only male athletes, we report relationships for both sportsmen and sportswomen. As predicted by evolutionary arguments, we found that male biathletes were judged more attractive if (unknown to the raters) they had achieved a higher peak performance (World Cup points score) in their career, whereas there was no significant relationship for female biathletes. Our findings show that elite male athletes display visible, attractive cues that reliably reflect their athletic performance.

Keywords: sexual signaling, whole-organism performance, endurance, evolutionary sports science, fWHR, mouth curvature

INTRODUCTION

The evolution of mating preferences for indicators of direct or indirect fitness benefits is fundamental to all major theories of sexual selection (Kokko et al. 2006; Kuijper et al. 2012). Although most research has focused on preferences for morphological ‘ornaments’ such as enlarged appendages or bright color patches (Andersson 1994; Andersson and Simmons 2006), evidence suggests that mating patterns are also influenced by behavioral and physiological characteristics, through their effects on whole-organism performance (Lailvaux and Irschick 2006; Husak and Fox 2008; Lailvaux and Husak 2014). Individual variation in performance can influence both intrasexual and intersexual interactions. In some animals, athletic ability (e.g. endurance, sprint speed) predicts the outcome of intrasexual competition, which in turn determines access to mating opportunities (e.g. beetles, crustaceans and lizards; reviewed in Lailvaux and Irschick 2006). In others, courtship behavior directed towards the opposite sex involves active displays of maximum power output, motor skill or stamina and these performance measures are associated with higher mating success (e.g. Anna’s hummingbirds, *Calypte anna*, Clark 2009; golden-collared manakins, *Manacus vitellinus*, Barske et al. 2011; Cuban burrowing cockroaches, *Byrsotria fumigata*, Mowles and Jepson 2015).

Competitive sport offers a unique setting in which to examine some of these issues in our own species. Recent studies on a range of different sports have suggested that women are attracted to men with higher sporting ability, based purely on static images of their face and upper shoulders. When shown facial photographs of elite sportsmen, women gave higher attractiveness ratings to National Football League quarterbacks with better passer ratings (Williams et al. 2010), cyclists who achieved a higher finishing position in the 2012 Tour de France (Postma 2014) and mixed martial artists who had won their bouts (Little et al. 2015). A study on soccer and ice hockey (Park et al. 2007) also reported higher attractiveness ratings for men who play in arguably more athletically demanding positions (strikers, goalkeepers/goalies) than those in other positions (defenders/defensemen), although detailed analysis of the workload in different soccer positions suggests a more complex picture (Bloomfield et al. 2007; Gil et al. 2007). While facial attractiveness is unlikely to have a direct impact on success in any of these sports, it has been suggested that facial cues to sporting performance could arise through multiple effects of testosterone and other androgens (Williams et al. 2010; Tsujimura and Banissy 2013; Zilioli et al. 2015). Androgens have been linked both to the development of facial structure during puberty (Weston et al. 2007) and to behavior in competitive interactions (Eisenegger et al. 2011; Oliveira and Oliveira 2014), though direct evidence for a common mechanism is weak at best (Bird et al. 2016).

According to evolutionary arguments, a female preference for more athletic men was selectively favored in our recent evolutionary past because pairing with such men offered direct or indirect benefits (Williams et al. 2010; Postma 2014; Longman et al. 2015). Such arguments are perhaps most relevant for endurance, i.e. sustained activity over long distances, which may have been an important determinant of foraging (hunting or scavenging) success in ancestral environments and for which humans have an unusual capacity among mammals (Carrier 1984; Bramble and Lieberman 2004; Lieberman and Bramble 2007). However, the extension of this evolutionary logic to performance in elite sports, and the empirical evidence proposed to support it, is hotly debated. Critics have argued that the reported effect sizes are weak, that findings from homogeneous groups of elite athletes cannot be generalized to the wider human population and that available performance metrics reflect variation in sport-specific training rather than biological indicators of heritable fitness (Smoliga and Zavorsky 2015, 2016; see counter-arguments in Postma 2016). Although the genetic basis of variation in elite athletic performance is disputed (Smoliga and Zavorsky 2016; Postma 2016), this debate overlooks a crucial point: a preference for more athletic males could evolve even if athleticism is not heritable. Indeed, one general conclusion from models of sexual selection is that preferences for direct (i.e. non-genetic) benefits typically evolve more easily than those for indirect (i.e. genetic) benefits (Kokko et al. 2006; Kuijper et al. 2012).

There are, however, other important limitations of much of the published research on attractiveness and sporting ability. First, performance in team sports (e.g. American football, soccer and ice hockey) is strongly dependent on the behavior of other individuals (i.e. the focal individual's team-mates). Even the Tour de France, which superficially may seem like an individual sport, has a well-known strategic, team-based element (Torgler 2007) that partly determines finishing positions in a given year. Although it seems likely that individual performance capacity would have partly contributed to the measured outcomes in these studies, a purer measure of athletic performance could be obtained by using an individual-level sport in which there is no team element.

A second limitation, specific to Postma's (2014) Tour de France study, is that attractiveness ratings may have been influenced by the raters' knowledge of the research aims. The online advertisement recruiting participants for this study explicitly stated that the aim was to investigate "the relationship between looks and performance" using "the portraits of professional cyclists that have taken part in the 2012 Tour de France" (Postma 2012). It is possible, therefore, that the reported relationship could have been driven by demand characteristics (Orne 1962) leading participants to associate more athletic-looking faces with higher attractiveness. To demonstrate a valid preference for more athletic individuals that is not driven by demand characteristics, it is important that explicit information about the sporting context is hidden from raters.

Finally, all previous studies have focused entirely on the relationship between facial attractiveness and sporting performance in male athletes, ignoring whether a similar relationship exists for female athletes. If the evolutionary explanation for this relationship is credible—that an ancestral preference for more athletic mates led to direct or indirect fitness benefits—then there are reasons to expect that the relationship will be different for females. Evidence suggests that in our recent evolutionary past, it was primarily men rather than women who engaged in hunting activities (Hawkes and Bliege Bird 2002; Marlowe 2007); the potential benefits for a man choosing a more athletic partner are less clear. In addition, the proposed role of testosterone as a mechanistic link between facial characteristics and athletic performance is more plausible for men than for women, given that the sexual divergence of human facial structure (Weston et al. 2007) and neuromuscular performance (Beunen and Malina 1988) coincides with a pubertal surge in testosterone production in men (Verdonck et al. 1999). For these reasons, we would expect the relationship between facial attractiveness and sporting performance to be weaker or even non-existent in women, compared to men. Examining the relationships for both sexes would therefore allow a more comprehensive test of evolutionary predictions.

Here we report a study that addresses all the above limitations. For the first time, we determine the relationship between facial attractiveness and sporting performance in both male and female athletes in an individual-based sport without any team element, using attractiveness judgements made by raters who were unaware of the sporting connection. We focus on the biathlon, a cross-country skiing race interspersed with rounds of target shooting that tests elements of both endurance and skill. Cross-country skiing requires a large amount of aerobic power, muscle strength (Neumayr et al. 2003), balance (Müller et al. 2011), coordination and endurance (Stöggl et al. 2010), while shooting requires the ability to compose oneself via breathing techniques so that the physiological demands of the skiing do not affect shooting accuracy (Sattlecker et al. 2007). The International Biathlon Union organizes an annual series of World Cup events in which men compete over distances of 10–20 km and women over 7.5–15 km, generating individual performance metrics each year for the top international competitors of both sexes. Independently, we obtained opposite-sex attractiveness ratings for facial photographs of World Cup biathletes from a sample of participants in the UK, where biathlon is not widely followed and therefore we could be confident that the ratings were not influenced by a perceived connection to sport.

METHODS

Athletes

We obtained data on all 173 athletes (89 men aged 19–38 years; 84 women aged 22–40 years) who competed in the biathlon at the 2014 Winter Olympic Games in Sochi, Russia. Passport-style photographs were downloaded from the Russian sports website Р-Спорт (R-Sport; archived at http://sochi2014.arch.articul.ru/www.sochi2014.com/en/biathlon-athletes.htm) and rescaled to a standard size (144 × 80 pixels). We discarded 23 photos (12 male, 11 female) that were of poor quality, or had features that potentially identified the subject as an athlete (e.g. national sports kit), or for whom performance data (see below) were unavailable. This left us with a sample of 78 male and 78 female photos, depicting only the head, neck and upper shoulders of the athlete, evenly lit against a plain background and directly facing the camera. We took two sets of measurements from these photos that previous research suggests may influence ratings of attractiveness and dominance: facial width-to-height ratio (fWHR; Fig. 1a), calculated as the bizygomatic width (distance between left and right cheekbones at the widest part of the face) divided by the upper facial height (distance between upper lip and brow) (Weston et al. 2007; Carré and McCormick 2008); and mouth curvature (Fig. 1b), calculated as the upturn of the mouth (vertical distance from mouth center to left and right corners) divided by the mouth width (distance between left and right corners) (Tamalas et al. 2016). We obtained the date of birth, height and weight for all of these athletes from the Р-Спорт website (see above).

To assess performance we used the 'World Cup total score' as defined by the International Biathlon Union (2016; section 15.8.4.1). This total, recalculated each season, comprises the points scored in all individual, non-relay World Cup events ('individual', 'sprint', 'pursuit' and 'mass start'), minus the two lowest scores; note that team-based events ('relay' and 'mixed relay') are excluded. The scoring system awards 60 points for winning a race and gradually decreasing points down to 40th place (for full details see International Biathlon Union 2016). We recorded each athlete's World Cup total score in every season from 2001–02 to 2013–14 inclusive, as archived on the International Biathlon Union's Datacenter (http://biathlonresults.com) and another biathlon statistics website (http://www.realbiathlon.com), and then took the highest score for each athlete as a measure of their career-best performance.

Raters

To rate the attractiveness of the athletes we recruited 25 male and 25 female participants (mean age 21.3 years, range 17–58) via e-mail, social media and opportunity sampling around the University of Bristol campus; most were undergraduate students. This number of raters is comparable to several previous studies using facial attractiveness judgements (e.g. n = 21 in Penton-Voak et al. 2001; n = 28 in Penton-Voak and Chang 2008; n = 30 in Williams et al. 2010; n = 33 in Little et al. 2015). Participation in the study was completely voluntary and no payment was offered.

Procedure

Participants were taken to a test room in the University of Bristol's Life Sciences Building, where they read and signed a consent form that provided basic information about the testing procedure (without revealing the study's aims or the connection to sport) and explained that they were free to withdraw at any stage. They then completed (at their own pace) a series of questions using keystrokes on a laptop computer, presented using E-Prime software version 2.0 (Psychology Software Tools 2002). After confirming their sex and age, the participants were shown the photos of opposite-sex biathletes in randomized order and asked to indicate (i) how physically attractive they found that person on a scale from 1 (very unattractive) to 7 (very attractive) and (ii) whether they recognized the person. At the end of the study they were asked to indicate their sexual orientation. All details of the procedure were approved by the University of Bristol Research Ethics Committee (ref. 12741).

Statistical analysis

One male rater identified himself as homosexual at the end of the task, so his ratings were omitted before analysis. We also omitted 76 cases (less than 2% of the sample; no more than three cases for any athlete) where the rater reported that they recognized the face, even though when probed by the experimenter none of these correctly identified that the faces belonged to elite athletes. This left us with a sample of n = 3,746 attractiveness scores for 78 male and 78 female biathletes, rated by 24 male and 25 female participants. Including all of the data (n = 3,900) did not change the patterns reported here (supplementary tables S3 and S4). The results were also the same when excluding the small number (six women and three men) of non-Caucasian biathletes (supplementary tables S5 and S6).

To analyze the factors affecting the variation in attractiveness ratings we ran a series of linear mixed-effects models (LMMs) using the packages *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2015) in R version 3.5.1 (R Core Team 2018). In all models, the athlete and rater identities were included as random effects to account for non-independent ratings. First, we fitted a model to the attractiveness data for both sexes combined, with fixed effects of athlete performance (highest World Cup total score), sex, age, height, body mass index (BMI = weight (kg) divided by height (m) squared) and a two-way interaction term between athlete sex and performance. BMI was used in place of weight to reduce problems with multicollinearity, given that weight and height measurements are very strongly correlated (44.0% shared variance between weight and height in female biathletes and 64.0% in male biathletes, compared to 3.4% and 1.7% respectively between BMI and height). Before analysis, the response variable (attractiveness rating) and all continuous predictors (age, height, BMI and performance) were converted to *Z* scores (i.e. standardized) within each sex by subtracting the mean and dividing by the standard deviation for that sex. We included both linear and quadratic terms for the effects of age, height and BMI. Because the sex × performance interaction term was significant, we then analyzed the data for each sex separately. Finally, we checked whether the observed relationships were mediated by mouth curvature or fWHR by including these measurements (also converted to *Z* scores) as additional predictors in the model.

The models were fitted using restricted maximum likelihood (REML) and the significance of fixed effects was assessed using Wald *t* tests with Satterthwaite-approximated degrees of freedom. Where significant effects were found we used likelihood-ratio tests (based on maximum likelihood (ML) estimation) to check whether the inclusion of random slopes (varying with rater identity) improved the fit of the model. Residual plots confirmed assumptions of normality and homoscedasticity for all models.

The full data set and R code are available as supplementary information archived in the Dryad digital repository (Fawcett et al. in press).

RESULTS

Raters varied significantly in the mean attractiveness rating they gave (random effect of rater identity, explaining 30.5% of the variation in ratings; LMM: $\chi^2_1 = 1820.9$, $P < 0.001$).

Despite this, there was significant variation among biathletes in their mean rated attractiveness (random effect of athlete identity, explaining 29.4% of variation; LMM: χ^2_1 = 1659.3, P < 0.001) and the raters showed strong agreement overall in which biathletes they found attractive (intra-class correlation r = 0.838, based on variance components from one-way ANOVA).

A model for both sexes combined, controlling for age, height and body mass index (BMI), revealed that the relationship between attractiveness and sporting performance (career-best World Cup total score) differed significantly between male and female biathletes (sex × performance interaction term: P = 0.010; Table 1). There was also significant variation among individual raters in how their ratings were related to athlete performance (random slope term, explaining 0.5% of variation; χ^2_2 = 14.2, P = 0.001). To decompose the sex × performance interaction term, we subsequently analyzed the sexes separately (Table 2). Among female biathletes, attractiveness ratings declined significantly with age, but there was no effect of performance (Table 2a, Fig. 2a). By contrast, male biathletes who had achieved a higher World Cup total score in their career were rated as significantly more attractive (Table 2b, Fig. 2b). All quadratic terms were non-significant (supplementary table S1), so were omitted from the final models shown here. This pattern of results matches evolutionary predictions, suggesting that women are sensitive to cues that reliably indicate athletic ability in men.

Previous work (Williams et al. 2010; Tsujimura and Banissy 2013; Zilioli et al. 2015) has suggested that sporting performance might covary with differences in facial structure linked to androgens. For our data set, however, although fWHR was positively related to sporting performance in male biathletes (linear regression: b ± s.e. = 0.236 ± 0.113, t_{70} = 2.08, P = 0.041), this morphological measure did not predict their facial attractiveness ratings (LMM: b ± s.e. = 0.023 ± 0.066, $t_{66.2}$ = 0.35, P = 0.730). Another possibility is that athletic ability is revealed not by facial structure but by facial expression, reflecting an athlete's confidence or past success. We found that mouth curvature (a proxy for smiling; Tamalas et al. 2016) was negatively related to sporting performance in male biathletes (linear regression: b ± s.e. = −0.319 ± 0.122, t_{70} = −2.62, P = 0.011), but again did not predict their facial attractiveness ratings (LMM: b ± s.e. = 0.067 ± 0.072, $t_{65.9}$ = 0.94, P = 0.353). Importantly, including fWHR and mouth curvature in our earlier models did not alter the pattern of other effects: as before, facial attractiveness was positively related to performance in male (LMM: b ± s.e. = 0.160 ± 0.071, $t_{72.4}$ = 2.27, P = 0.026), but not female (LMM: b ± s.e. = 0.011 ± 0.077, $t_{66.0}$ = 0.14, P = 0.888), biathletes (supplementary table S2).

DISCUSSION

Our analysis shows that male biathletes who had achieved a higher World Cup total score in their career were judged as more attractive by the opposite sex based solely on a photograph of their face and upper shoulders, whereas there was no such relationship for female biathletes. These patterns hold when controlling for age, height and BMI. Previous studies have shown that attractiveness ratings are higher for elite sportsmen who won their last mixed martial arts bout (Little et al. 2015) or achieved a higher finishing position in the 2012 Tour de France cycling race (Postma 2014), while ours shows that attractiveness is also linked to career-best performance in an annual competition in which individuals are 'playing the field', without any dyadic or strategic team-based element. Most importantly, our study is the first to examine this relationship in both sexes and show that it only exists for male athletes. By keeping the sports connection hidden from raters, we ensured that the observed relationships could not be driven by demand characteristics. Our results therefore provide strong evidence that photographs of successful male athletes contain cues that are attractive to the opposite sex.

There are at least four possible explanations for these results. The first possibility is that, as suggested by some evolutionary hypotheses based on intersexual selection, more athletic men have physical characteristics that reliably signal their greater performance capacity, and women are attuned to those characteristics because in ancestral environments they predicted direct or indirect fitness benefits (Williams et al. 2010; Postma 2014; Longman et al. 2015). For this hypothesis to work requires that athletes varying in their performance measures show perceptible differences in features of the head (including face), neck or upper shoulders in static photographs, given that this was the only information seen by our participants. As a candidate cue we examined fWHR, a sexually dimorphic measure possibly linked to hormonal changes during puberty (Verdonck et al. 1999; Carré and McCormick 2008) and correlated with aggressive behavior (Carré and McCormick 2008; Carré et al. 2009), the outcome of violent conflicts (Zilioli 2015; Stirrat et al. 2012) and sporting success (Tsujimura and Banissy; but see Mayew 2013). A meta-analysis of studies investigating fWHR concluded that it influences ratings of dominance or threat and, to a lesser extent, ratings of attractiveness (Geniole et al. 2015). In our data set, fWHR was positively related to peak performance in male biathletes but not to their rated attractiveness, and including it as a predictor in our statistical models did not explain the observed relationship between performance and attractiveness. There may well be other cues besides fWHR in the face, neck or upper shoulders that are consistently related to athletic performance; further work using more detailed morphometric comparisons would be needed to identify what these cues might be.

A second possible explanation is that success in World Cup events is reflected in an athlete’s facial expression, which in turn influences their attractiveness to the opposite sex. For example, athletes who perform better than their rivals may be happier or more confident, either as a direct result of their success (e.g. good performances lead to higher confidence and more positive mood states) or because pre-existing differences in confidence have an important influence on outcomes in elite sport (Moritz et al. 2000; Feltz 2007; Hays et al. 2009), perhaps particularly in men (Woodman and Hardy 2003). To investigate this possibility we quantified mouth curvature, a measure of facial expression indicative of smiling (Tamalas et al. 2016). Previous research suggests that smiling can enhance attractiveness (Jones et al. 2006; Golle et al. 2014), but perhaps only in women, with a neutral (Penton-Voak and Chang 2008) or even negative (Tracy and Beall 2011) effect of smiling on male attractiveness. In our study, including mouth curvature as an additional predictor did not account for the observed relationship between performance and attractiveness in men, despite a significant negative relationship between mouth curvature and performance. Future work analyzing a more extensive set of feature point coordinates (Benson and Perrett 1991; Tiddeman et al. 2001) may reveal subtler differences in facial expression that potentially influence attractiveness judgements.

A third possibility is that athletes who are judged more facially attractive receive more support and investment from an early age, ultimately leading to an improved career performance compared to less attractive athletes. Studies suggest that attractive people are treated more favorably than less attractive people in a range of contexts, leading to better economic prospects, a greater chance of being hired for jobs and even more affectionate interactions with their mothers (Langlois et al. 2000; Little 2014). Such advantages could extend into the sporting domain if, for example, better-looking athletes are more likely to be selected for high-performance programs, receive extra attention from coaching staff and secure lucrative sponsorship deals, potentially enhancing their career performance. While intriguing, we consider this to be an unlikely explanation for our results, because if anything it would predict that the positive relationship between sporting performance and facial attractiveness should be stronger in female than male athletes. Sports coaching is dominated by men (Knoppers 1992; Walker and Bopp 2011) and much has been written about the power of male coaches over their athletes (Brackenridge 1997; Fasting and Brackenridge 2009), particularly the circumstances under which this power can be exploited and lead to sexual harassment or abuse of female athletes (Cense and Brackenridge 2001; Nielsen 2001; Fasting et al. 2003, 2004). Furthermore, while biased investment in more attractive athletes may have a strong influence on progression to elite level, the impact on performance outcomes among those who have successfully made it to that level is likely to be much weaker. Nonetheless, investigating attractiveness biases in sport would be a valuable direction for future work. Evidence from the German Bundesliga suggests that a footballer's market value is enhanced by his facial attractiveness, independent of actual performance ratings (Rosar et al. 2017), but to our knowledge no studies have addressed whether coaching behavior and other aspects of athlete development are affected by physical attractiveness, in either sex.

A final possibility is that more successful athletes spend more time, effort and money enhancing their attractiveness through personal grooming, cosmetic surgery or other means. We were unable to control for the use of make-up in the images, although it is important to note that these were fairly standardized, passport-style photographs rather than publicity shots. While this explanation could potentially apply to some higher-profile sports in which success generates fame, with accompanying publicity and advertising deals, it seems unlikely to explain our results here, particularly given the absence of an effect in women. Nonetheless, future studies could improve on our methodology by ensuring greater standardization of the photos (e.g. covering of hair, no make-up).

Our study complements related findings in Tour de France cyclists (Postma 2014) and mixed martial artists (Little et al. 2015) and adds to the nascent field of evolutionary sports science (Wilson et al. 2017), highlighting the value of sports data as a rich resource for investigating how selection acts on psychological and physiological aspects of athletic performance. Using annual performance measures from the biathlon World Cup, we found that male, but not female, biathletes who had achieved a higher career peak were rated as more physically attractive by the opposite sex. This pattern is consistent with the evolutionary hypothesis that a female preference for more athletic men evolved through sexual selection, but also with other potential explanations. Further work is required to identify the specific cues that make better male athletes more attractive and to establish whether those cues directly reveal natural variation in sporting ability, confidence arising from differential success or biased investment in their athletic development.

REFERENCES

Andersson M. 1994. Sexual selection. Princeton (NJ): Princeton University Press.

Andersson M, Simmons LW. 2006. Sexual selection and mate choice. Trends Ecol Evol. 21:296–302.

Barske J, Schlinger BA, Wikelski M, Fusani L. 2011. Female choice for male motor skills. Proc R Soc B. 278:3523–3528.

Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw. 67:1–48.

Benson PJ, Perrett DI. 1991. Synthesising continuous-tone caricatures. Image Vis Comput. 9:123–129.

Beunen G, Malina R. 1988 Growth and physical performance relative to the timing of the adolescent spurt. Exerc Sport Sci Rev. 16:503–540.

Bird BM, Cid Jofré VS, Geniole SN, Welker KM, Zilioli S, Maestripieri D, Arnocky S, Carré JM. 2016. Does the facial width-to-height ratio map onto variability in men's testosterone concentrations? Evol Hum Behav. 37:392–398.

Bloomfield J, Polman R, O'Donoghue P. 2007. Physical demands of different positions in FA Premier League soccer. J Sports Sci Med. 6:63–70.

Brackenridge C. 1997. 'He owned me basically…'. Women's experience of sexual abuse in sport. Int Rev Sociol Sport. 32:115–130.

Bramble DM, Lieberman DE. 2004. Endurance running and the evolution of *Homo*. Nature. 432:345–352.

Carré JM, McCormick CM. 2008. In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. Proc R Soc B. 275:2651–2656.

Carré JM, McCormick CM, Mondloch CJ. 2009. Facial structure is a reliable cue of aggressive behavior. Psychol Sci. 20:1194–1198.

Carrier DR. 1984. The energetic paradox of human running and hominid evolution. Curr Anthropol. 25:483–495.

Cense M, Brackenridge C. 2001. Temporal and developmental risk factors for sexual harassment and abuse in sport. Eur Phy Educ Rev. 7:61–79.

Clark CJ. 2009. Courtship dives of Anna’s hummingbird offer insights into flight performance limits. Proc R Soc B. 276:3047–3052.

Eisenegger C, Haushofer J, Fehr E. 2011. The role of testosterone in social interaction. Trends Cogn Sci. 15:263–271.

Fasting K, Brackenridge C. 2009. Coaches, sexual harassment and education. Sport Educ Soc. 14:21–35.

Fasting K, Brackenridge C, Sundgot-Borgen J. 2003. Experiences of sexual harassment and abuse among Norwegian elite female athletes and nonathletes. Res Q Exerc Sport. 74:84–97.

Fasting K, Brackenridge C, Sundgot-Borgen J. 2004. Prevalence of sexual harassment among Norwegian female elite athletes in relation to sport type. Int Rev Sociol Sport. 39:373–386.

Fawcett TW, Ewans J, Lawrence A, Radford AN. In press. Data and R code from: Attractiveness is positively related to World Cup performance in male, but not female, biathletes. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.nk764n1.

Feltz DL. 2007. Self-confidence and sports performance. In: Smith D, Bar-Eli M, editors. Essential readings in sport and exercise psychology. Champaign (IL): Human Kinetics. p. 278–294.

Geniole SN, Denson TF, Dixson BJ, Carré JM, McCormick CM. 2015. Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. PLoS ONE. 10:e0132726.

Gil SM, Gil J, Ruiz F, Irazusta A, Irazusta J. 2007. Physiological and anthropometric characteristics of young soccer players according to their playing position: relevance for the selection process. J Strength Cond Res. 21:438–445.

Golle J, Mast FW, Lobmaier JS. 2014. Something to smile about: the interrelationship between attractiveness and emotional expression. Cogn Emot. 28:298–310.

Hawkes K, Bliege Bird R. 2002. Showing off, handicap signaling, and the evolution of men's work. Evol Anthropol. 11:58–67.

Hays K, Thomas O, Maynard I, Bawden M. 2009. The role of confidence in world-class sport performance. J Sports Sci. 27:1185–1199.

Husak JF, Fox SF. 2008. Sexual selection on locomotor performance. Evol Ecol Res. 10:213–228.

International Biathlon Union. 2016. IBU biathlon guide 2016/2017. Salzburg, Austria: International Biathlon Union. Available from: http://www.biathlonworld.com/downloads/.

Jones BC, DeBruine LM, Little AC, Conway CA, Feinberg DR. 2006. Integrating gaze direction and expression in preferences for attractive faces. Psychol Sci. 17:588–591.

Knoppers A. 1992. Explaining male dominance and sex segregation in coaching: three approaches. Quest. 44:210–227.

Kokko H, Jennions MD, Brooks R. 2006. Unifying and testing models of sexual selection. Annu Rev Ecol Evol Syst. 37:43–66.

Kuijper B, Pen I, Weissing FJ. 2012. A guide to sexual selection theory. Annu Rev Ecol Evol Syst. 43:287–311.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2015. lmerTest: tests in linear mixed effects models (version 2.0-33). Available from: http://CRAN.R-project.org/package=lmerTest.

Lailvaux SP, Husak JF. 2014. The life history of whole-organism performance. Q Rev Biol. 89:285–318.

Lailvaux SP, Irschick DJ. 2006. A functional perspective on sexual selection: insights and future prospects. Anim Behav. 72:263–273.

Langlois JH, Kalanakis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. Psychol Bull. 126:390–423.

Lieberman DE, Bramble DM. 2007. The evolution of marathon running. Sports Med. 37:288–290.

Little AC. 2014. Facial attractiveness. WIREs Cogn Sci. 5:621–634.

Little AC, Třebický V, Havlíček J, Roberts SC, Kleisner K. 2015. Human perception of fighting ability: facial cues predict winners and losers in mixed martial arts fights. Behav Ecol. 26:1470–1475.

Longman D, Wells JCK, Stock JT. 2015. Can persistence hunting signal male quality? A test considering digit ratio in endurance athletes. PLoS ONE. 10:e0121560.

Marlowe FW. 2007. Hunting and gathering: the human sexual division of foraging labor. Cross Cult Res. 41:170–195.

Mayew WJ. 2013. Reassessing the association between facial structure and baseball performance: a comment on Tsujimura and Banissy (2013). Biol Lett. 9:20130538.

Moritz SE, Feltz DL, Fahrbach KR, Mack DE. 2000. The relation of self-efficacy measures to sport performance: a meta-analytic review. Res Q Exerc Sport. 71:280–294.

Mowles SL, Jepson NM. 2015. Physiological costs of repetitive courtship displays in cockroaches handicap locomotor performance. PLoS ONE. 10:e0143664.

Müller E, Gimpl M, Kirchner S, Kroll J, Jahnel R, Niebauer J, Niederseer D, Scheiber P. 2011. Salzburg Skiing for the Elderly Study: influence of alpine skiing on aerobic capacity, strength, power, and balance. Scand J Med Sci Sports. 21:9–22.

Neumayr G, Hoertnagl H, Pfister R, Koller A, Eibl G, Raas E. 2003. Physical and physiological factors associated with success in professional alpine skiing. Int J Sports Med. 24:571–575.

Nielsen JT. 2001. The forbidden zone: intimacy, sexual relations and misconduct in the relationship between coaches and athletes. Int Rev Sociol Sport. 36:165–182.

Oliveira GA, Oliveira RF. 2014. Androgen responsiveness to competition in humans: the role of cognitive variables. Neurosci Neuroecon. 3:19–32.

Orne MT. 1962. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. Am Psychol. 17:776–783.

Park JH, Buunk AP, Weiling MB. 2007. Does the face reveal athletic flair? Positions in team sports and facial attractiveness. Pers Individ Dif. 43:1960–1965.

Penton-Voak IS, Chang HY. 2008. Attractiveness judgements of individuals vary across emotional expression and movement conditions. J Evol Psychol. 6:89–100.

Penton-Voak IS, Jones BC, Little AC, Baker S, Tiddeman B, Burt DM, Perrett DI. Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. Proc R Soc B. 268:1617–1623.

Postma E. 2012. Human sexual selection survey: Do nice guys finish first, or are looks deceiving? (Accessed 22 August 2012). Archived at http://life.mcmaster.ca/~brian/evoldir/Archive/Mnth_Review_Aug_12.pdf.

Postma E. 2014. A relationship between attractiveness and performance in professional cyclists. Biol Lett. 10:20130966.

Postma E. 2016. Why we should *not* dismiss a relationship between attractiveness and performance: a comment on Smoliga and Zavorsky (2015). Biol Lett. 12:20160068.

Psychology Software Tools. 2002. E-Prime version 2.0. Available from: http://www.pstnet.com.

R Core Team. 2018. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Rosar U, Hagenah J, Klein M. 2017. Physical attractiveness and monetary success in German Bundesliga. Soccer Soc. 18:102–120.

Sattlecker G, Müller E, Lindinger S. 2007. Performance determining factors in biathlon shooting. In: V Linnamo PV Komi, E Müller, editors. Science and Nordic Skiing. Oxford (UK): Meyer and Meyer Sport. p. 257–265.

Smoliga JM, Zavorsky GS. 2015. Faces and fitness: attractive evolutionary relationship or ugly hypothesis? Biol Lett. 11:20150839.

Smoliga JM, Zavorsky GS. 2016. Exercise physiology and sports science must be considered in evolutionary theories regarding human performance: a reply to Postma (2016). Biol Lett. 12:20160856.

Stirrat M, Stulp G, Pollet TV. 2012. Male facial width is associated with death by contact violence: narrow-faced males are more likely to die from contact violence. Evol Hum Behav. 33:551–556.

Stöggl T, Enqvist J, Müller E, Holmberg H-C. 2010. Relationships between body composition, body dimensions, and peak speed in cross-country sprint skiing. J Sports Sci. 28:161–169.

Tamalas SN, Mavor KI, Axelsson J, Sundelin T, Perrett DI. 2016. Eyelid-openness and mouth curvature influence perceived intelligence beyond attractiveness. J Exp Psychol Gen. 145:603–620.

Tiddeman BP, Burt DM, Perrett DI. 2001. Prototyping and transforming facial textures for perception research. IEEE Comput Graph Appl. 21:42–50.

Torgler B. 2007. “La Grande Boucle”: determinants of success at the Tour de France. J Sports Econ. 8:317–331.

Tracy JL, Beall AT. 2011. Happy guys finish last: the impact of emotional expressions on sexual attraction. Emotion. 11:1379–1387.

Tsujimura H, Banissy MJ. 2013. Human face structure correlates with professional baseball performance: insights from professional Japanese baseball players. Biol Lett. 9:20130140.

Verdonck A, Gaethofs M, Carels C, de Zegher F. 1999. Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. Eur J Orthod. 21:137–143.

Walker NA, Bopp T. 2011. The underrepresentation of women in the male-dominated sport workplace: perspectives of female coaches. J Workpl Rights 15:47–64.

Weston EM, Friday AE, Liò P. 2007. Biometric evidence that sexual selection has shaped the hominin face. PLoS ONE. 2:e710.

Williams KM, Park JH, Weiling MB. 2010. The face reveals athletic flair: better National Football League quarterbacks are better looking. Pers Individ Dif. 48:112–116.

Wilson RS, David GK, Murphy SC, Angilletta MJ Jr, Niehaus AC, Hunter AH, Smith MD. 2017. Skill not athleticism predicts individual variation in match performance of soccer players. Proc R Soc B. 284:20170953.

Woodman T, Hardy L. 2003. The relative impact of cognitive anxiety and self-confidence upon sport performance: a meta-analysis. J Sports Sci. 21:443–457.

Zilioli S, Sell AN, Stirrat M, Jagore J, Vickerman W, Watson NV. 2015. Face of a fighter: bizygomatic width as a cue of formidability. Aggress Behav. 41:322–330.

Figure 1. Measurement of (a) facial width-to-height ratio (fWHR) and (b) mouth curvature from portrait photographs. We calculated fWHR as W/H (following Weston et al. 2007), where W is the bizygomatic width (distance between left and right cheekbones at the widest part of the face) and H is the upper facial height (distance between upper lip and brow). We calculated mouth curvature as Y/X (following Tamalas et al. 2016), where Y is the upturn of the mouth (vertical distance from mouth center to left and right corners) and X is the mouth width (distance between left and right corners). Note that this image does not depict one of the biathletes used in this study, but is shown purely for illustrative purposes.

Figure 2. Mean standardized attractiveness of (a) female and (b) male biathletes as rated by the opposite sex, in relation to their career-best performance (highest World Cup total score). Dots represent individual athletes. The thick black line in panel (b) shows the significant (P = 0.017) positive relationship between performance and attractiveness in male biathletes from a linear mixed-effects model controlling for age, height and body mass index, with random intercepts for athlete and rater identity and a random slope term (varying among raters) for the effect of performance (rater-specific relationships shown as thin grey lines). The corresponding relationship was non-significant (P = 0.933) for female biathletes.

Figure 1

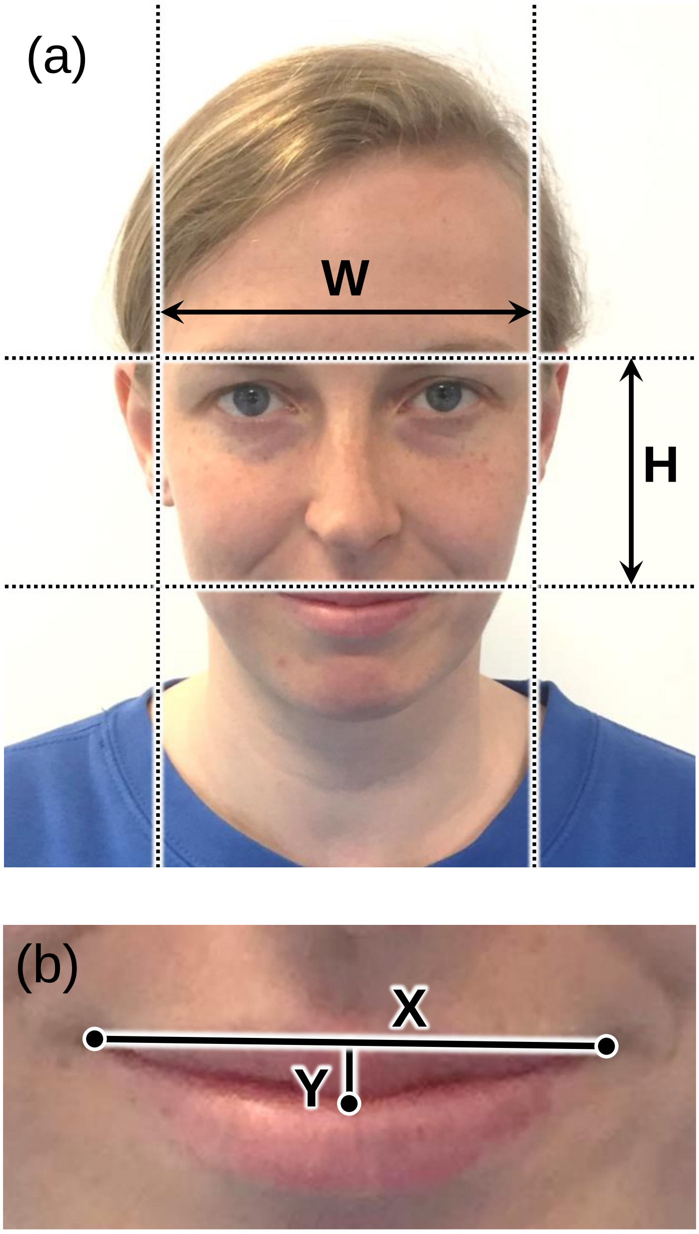

Figure 2

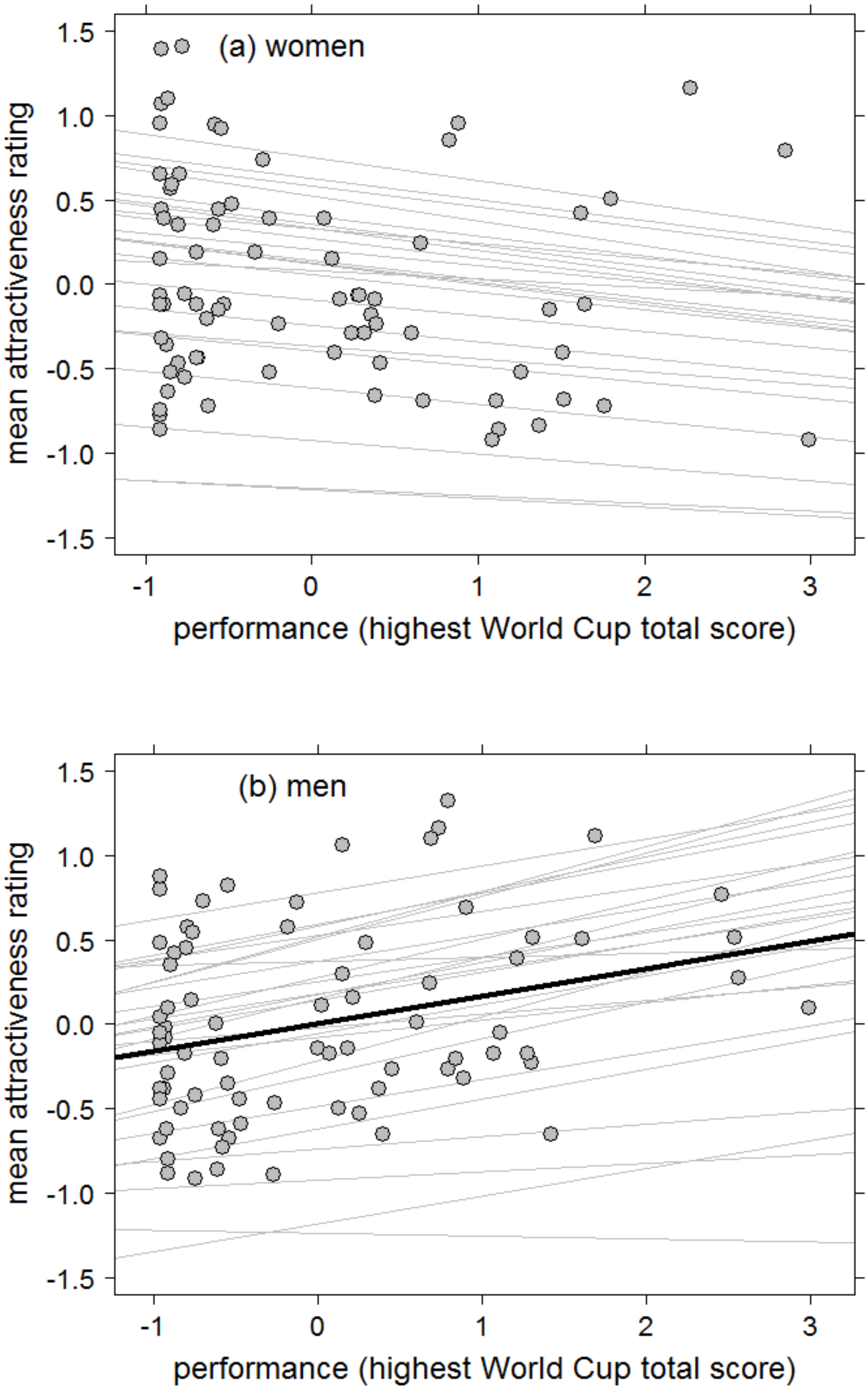
(a) women
mean attractiveness rating
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-1
0
1
2
3
performance (highest World Cup total score)
(b) men
mean attractiveness rating
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-1
0
1
2
3
performance (highest World Cup total score)

Table 1. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings for biathletes (both sexes, n = 156).

Fixed effect	Estimate ± SE	t	d.f.*	P
intercept	0.006 ± 0.129	0.05	76.5	0.964
sex (male)	−0.014 ± 0.182	−0.08	77.2	0.938
age	−0.085 ± 0.049	−1.74	148.9	0.084
height	0.011 ± 0.047	0.23	148.8	0.820
BMI	0.043 ± 0.046	0.93	148.8	0.353
performance†	−0.070 ± 0.070	−1.00	159.0	0.319
sex × performance	**0.245 ± 0.094**	**2.61**	**159.3**	**0.010**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters (χ^2_2 = 14.2, P = 0.001)

Table 2. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings separately for (a) female and (b) male biathletes.

Fixed effect	Estimate ± SE	t	d.f.*	P
(a) women (n = 78)				
intercept	0.006 ± 0.129	0.05	38.9	0.964
age†	**−0.209 ± 0.078**	**−2.69**	**76.2**	**0.009**
height	0.016 ± 0.070	0.24	73.0	0.814
BMI	−0.060 ± 0.069	−0.87	72.9	0.388
performance	−0.007 ± 0.077	−0.09	73.0	0.933
(b) men (n = 78)				
intercept	−0.008 ± 0.127	−0.07	37.3	0.947
age	−0.017 ± 0.063	−0.27	73.0	0.789
height	0.000 ± 0.062	0.01	72.9	0.995
BMI	0.109 ± 0.062	1.77	72.9	0.081
performance‡	**0.159 ± 0.065**	**2.45**	**81.3**	**0.017**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters (χ^2_2 = 10.6, P = 0.005)
‡slope varies significantly among raters (χ^2_2 = 16.3, P < 0.001)

SUPPLEMENTARY INFORMATION

Table S1. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings for biathletes (both sexes, n = 156), including linear and quadratic terms for continuous predictors.

Fixed effect	Estimate ± SE	t	d.f.*	P
intercept	0.135 ± 0.142	0.96	102.5	0.342
sex (male)	−0.014 ± 0.181	−0.08	76.7	0.938
age				
linear	−0.049 ± 0.054	−0.92	145.9	0.361
quadratic	−0.052 ± 0.040	−1.28	146.1	0.201
height				
linear	0.012 ± 0.047	0.25	145.8	0.803
quadratic	−0.053 ± 0.033	−1.60	145.8	0.111
BMI				
linear	0.065 ± 0.048	1.38	145.8	0.171
quadratic	−0.026 ± 0.031	−0.84	145.7	0.404
performance†	−0.092 ± 0.071	−1.30	155.8	0.197
sex × performance	**0.265 ± 0.094**	**2.81**	**156.1**	**0.006**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters (χ^2_2 = 14.1, P = 0.001)

Table S2. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings separately for (a) male and (b) female biathletes, controlling for mouth curvature and facial width-to-height ratio (fWHR).

Fixed effect	Estimate ± SE	*t*	d.f.	*P*
(a) women (*n* = 73)				
intercept	0.051 ± 0.130	0.40	39.3	0.697
age†	**−0.278 ± 0.081**	**−3.42**	**70.7**	**0.001**
height	−0.003 ± 0.071	−0.05	66.0	0.963
BMI	−0.108 ± 0.072	−1.50	65.9	0.139
mouth curvature	−0.003 ± 0.062	−0.06	66.0	0.956
fWHR	0.065 ± 0.068	0.95	65.9	0.345
performance	0.011 ± 0.077	0.14	66.0	0.888
(b) men (*n* = 73)				
intercept	0.003 ± 0.129	0.02	38.7	0.981
age	−0.022 ± 0.065	−0.33	66.0	0.739
height	0.005 ± 0.065	0.08	65.9	0.936
BMI	0.104 ± 0.063	1.64	65.9	0.107
mouth curvature	0.067 ± 0.072	0.94	65.9	0.353
fWHR	0.023 ± 0.066	0.35	66.2	0.730
performance‡	**0.160 ± 0.071**	**2.27**	**72.4**	**0.026**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters ($\chi^2_2 = 17.4$, $P < 0.001$)
‡slope varies significantly among raters ($\chi^2_2 = 13.6$, $P = 0.001$)

Table S3. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings for biathletes (both sexes, n = 156), including ratings from one homosexual rater and ratings where the rater reported that they recognized the face (3,900 ratings in total).

Fixed effect	Estimate ± SE	t	d.f.*	P
intercept	0.000 ± 0.125	0.00	81.6	> 0.999
sex (male)	0.000 ± 0.177	0.00	81.6	> 0.999
age	−0.088 ± 0.049	−1.80	149.0	0.074
height	0.009 ± 0.047	0.20	149.0	0.843
BMI	0.045 ± 0.046	0.98	149.0	0.329
performance†	−0.071 ± 0.070	−1.01	157.7	0.312
sex × performance	**0.246 ± 0.094**	**2.61**	**158.3**	**0.010**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters (χ^2_2 = 12.3, P = 0.002)

Table S4. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings separately for (a) female and (b) male biathletes, including ratings from one homosexual rater and ratings where the rater reported that they recognized the face (3,900 ratings in total).

Fixed effect	Estimate ± SE	*t*	d.f.*	*P*
(a) women (*n* = 78)				
intercept	0.000 ± 0.125	0.00	42.5	> 0.999
age†	**−0.207 ± 0.078**	**−2.66**	**75.9**	**0.010**
height	0.013 ± 0.070	0.18	73.0	0.858
BMI	−0.058 ± 0.069	−0.84	73.0	0.402
performance	−0.009 ± 0.077	−0.12	73.0	0.903
(b) men (*n* = 78)				
intercept	0.000 ± 0.124	0.00	38.3	> 0.999
age	−0.024 ± 0.063	−0.38	73.0	0.703
height	0.001 ± 0.062	0.02	73.0	0.982
BMI	0.114 ± 0.062	1.83	73.0	0.071
performance‡	**0.160 ± 0.065**	**2.46**	**81.0**	**0.016**

*denominator degrees of freedom derived using Satterthwaite approximation

†slope varies significantly among raters ($\chi^2_2 = 10.0$, $P = 0.007$)

‡slope varies significantly among raters ($\chi^2_2 = 16.1$, $P < 0.001$)

Table S5. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings for biathletes (both sexes, $n = 147$), excluding those biathletes identified as non-Caucasian (one Korean, one Chinese and one Japanese male biathlete, plus one Korean, two Japanese and three Chinese female biathletes).

Fixed effect	Estimate ± SE	*t*	d.f.*	*P*
intercept	0.061 ± 0.132	0.46	78.2	0.644
sex (male)	−0.072 ± 0.185	−0.39	78.0	0.698
age	−0.093 ± 0.050	−1.86	140.0	0.065
height	−0.010 ± 0.049	−0.20	139.9	0.845
BMI	0.036 ± 0.048	0.75	139.9	0.452
performance†	−0.106 ± 0.072	−1.47	151.2	0.142
sex × performance	**0.288 ± 0.096**	**3.00**	**151.6**	**0.003**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters ($\chi^2_2 = 18.3$, $P < 0.001$)

Table S6. Estimates of fixed effects in a linear mixed-effects model predicting opposite-sex attractiveness ratings separately for (a) female and (b) male biathletes, excluding those biathletes identified as non-Caucasian (one Korean, one Chinese and one Japanese male biathlete, plus one Korean, two Japanese and three Chinese female biathletes).

Fixed effect	Estimate ± SE	*t*	d.f.*	*P*
(a) women (*n* = 72)				
intercept	0.072 ± 0.131	0.55	37.9	0.588
age†	**−0.259 ± 0.078**	**−3.32**	**72.4**	**0.001**
height	−0.028 ± 0.070	−0.40	67.0	0.689
BMI	−0.109 ± 0.070	−1.57	66.9	0.121
performance	−0.021 ± 0.074	−0.28	67.0	0.782
(b) men (*n* = 75)				
intercept	−0.010 ± 0.128	−0.08	38.3	0.936
age	−0.008 ± 0.066	−0.11	70.0	0.910
height	0.003 ± 0.065	0.05	69.9	0.965
BMI	0.122 ± 0.065	1.87	69.9	0.066
performance‡	**0.159 ± 0.067**	**2.38**	**78.1**	**0.020**

*denominator degrees of freedom derived using Satterthwaite approximation
†slope varies significantly among raters ($\chi^2_2 = 18.2$, $P < 0.001$)
‡slope varies significantly among raters ($\chi^2_2 = 16.5$, $P < 0.001$)