

PERCEPTIONS OF PEOPLE: CUES TO UNDERLYING PHYSIOLOGY AND PSYCHOLOGY

EDITED BY: Kok Wei Tan, Lisa L. M. Welling, Ian D. Stephen, Alex L. Jones
and Danielle Sulikowski
PUBLISHED IN: Frontiers in Psychology





frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88963-748-5

DOI 10.3389/978-2-88963-748-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PERCEPTIONS OF PEOPLE: CUES TO UNDERLYING PHYSIOLOGY AND PSYCHOLOGY

Topic Editors:

Kok Wei Tan, University of Reading Malaysia, Malaysia

Lisa L. M. Welling, Oakland University, United States

Ian D. Stephen, Macquarie University, Australia

Alex L. Jones, Swansea University, United Kingdom

Danielle Sulikowski, Charles Sturt University, Australia

Citation: Tan, K. W., Welling, L. L. M., Stephen, I. D., Jones, A. L., Sulikowski, D., eds. (2020). Perceptions of People: Cues to Underlying Physiology and Psychology. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-748-5

Table of Contents

- 05 Editorial: Perceptions of People: Cues to Underlying Physiology and Psychology**
Danielle Sulikowski, Kok Wei Tan, Alex L. Jones, Lisa L. M. Welling and Ian D. Stephen
- 08 Evidence That the Hormonal Contraceptive Pill is Associated With Cosmetic Habits**
Carlota Batres, Aurélie Porcheron, Gwenaél Kaminski, Sandra Courrèges, Frédérique Morizot and Richard Russell
- 13 The Relationship Between Observers' Self-Attractiveness and Preference for Physical Dimorphism: A Meta-Analysis**
Lijun Chen, Xiaoliu Jiang, Huiyong Fan, Ying Yang and Zhihong Ren
- 27 The Relative Contribution of Jawbone and Cheekbone Prominence, Eyebrow Thickness, Eye Size, and Face Length to Evaluations of Facial Masculinity and Attractiveness: A Conjoint Data-Driven Approach**
Justin K. Mogilski and Lisa L. M. Welling
- 38 360 Degrees of Facial Perception: Congruence in Perception of Frontal Portrait, Profile, and Rotation Photographs**
Vít Třebický, Jitka Fialová, David Stella, Zuzana Štěrbová, Karel Kleisner and Jan Havlíček
- 49 Judging Others by Your Own Standards: Attractiveness of Primate Faces as Seen by Human Respondents**
Silvie Rádlová, Eva Landová and Daniel Frynta
- 65 Facial Adiposity, Attractiveness, and Health: A Review**
Stefan de Jager, Nicoleen Coetzee and Vinet Coetzee
- 81 The Influence of Body Composition Effects on Male Facial Masculinity and Attractiveness**
Xue Lei, Iris J. Holzleitner and David I. Perrett
- 95 Predictors of Fighting Ability Inferences Based on Faces**
Vít Třebický, Jitka Fialová, David Stella, Klára Coufalová, Radim Pavelka, Karel Kleisner, Radim Kuba, Zuzana Štěrbová and Jan Havlíček
- 106 Toward a New Approach to Cross-Cultural Distinctiveness and Typicality of Human Faces: The Cross-Group Typicality/Distinctiveness Metric**
Karel Kleisner, Šimon Pokorný and S. Adil Saribay
- 119 She Always Steps in the Same River: Similarity Among Long-Term Partners in Their Demographic, Physical, and Personality Characteristics**
Zuzana Štěrbová, Petr Tureček and Karel Kleisner
- 134 Do Single Men Smell and Look Different to Partnered Men?**
Mehmet K. Mahmut and Richard J. Stevenson
- 143 The Interplay Between Economic Status and Attractiveness, and the Importance of Attire in Mate Choice Judgments**
Amany Gouda-Vossos, Robert C. Brooks and Barnaby J. W. Dixon
- 157 Roar of a Champion: Loudness and Voice Pitch Predict Perceived Fighting Ability but Not Success in MMA Fighters**
Pavel Šebesta, Vít Třebický, Jitka Fialová and Jan Havlíček

- 166** *Evolutionary Theories and Men's Preferences for Women's Waist-to-Hip Ratio: Which Hypotheses Remain? A Systematic Review*
Jeanne Bovet
- 186** *Skin Color Preferences in a Malaysian Chinese Population*
Kok Wei Tan and Ian D. Stephen
- 192** *Vocal Parameters of Speech and Singing Covary and are Related to Vocal Attractiveness, Body Measures, and Sociosexuality: A Cross-Cultural Study*
Jaroslava Varella Valentova, Petr Tureček, Marco Antonio Corrêa Varella, Pavel Šebesta, Francisco Dyonisio C. Mendes, Kamila Janaina Pereira, Lydie Kubicová, Petra Stolařová and Jan Havlíček



Editorial: Perceptions of People: Cues to Underlying Physiology and Psychology

Danielle Sulikowski^{1*}, Kok Wei Tan², Alex L. Jones³, Lisa L. M. Welling⁴ and Ian D. Stephen^{5,6,7}

¹ Perception and Performance Research Group, School of Psychology, Charles Sturt University, Bathurst, NSW, Australia,

² School of Psychology and Clinical Language Sciences, University of Reading Malaysia, Gelang Patah, Malaysia,

³ Department of Psychology, College of Human and Health Sciences, Swansea University, Swansea, United Kingdom,

⁴ Psychology Department, Oakland University, Rochester, MI, United States, ⁵ Department of Psychology, Macquarie

University, Sydney, NSW, Australia, ⁶ Perception in Action Research Centre, Macquarie University, Sydney, NSW, Australia,

⁷ Body Image and Ingestive Behaviour Group, Macquarie University, Sydney, NSW, Australia

Keywords: mate quality signals, health, fertility, formidability, person perception

Editorial on the Research Topic

Perceptions of People: Cues to Underlying Physiology and Psychology

Our perceptual sensitivity to cues of socially and sexually relevant physiological and psychological traits in others is remarkable. For such sensitivity to evolve, the directly perceptible qualities of others (which include intrinsic physical traits, such as height, weight, body odor, facial morphology, and body shape; as well as behaviorally modified appearance cues, such as those produced by clothing and makeup; and vocal parameters) must afford at least somewhat accurate judgements of others' health, fertility, formidability, personality, or other fitness-relevant capacities. The current issue examines whether a variety of perceptible qualities present potentially valid cues of underlying physiology and psychology, and/or the extent to which such cues support adaptive judgements of others.

The human face is a highly complex signaling system, and not surprisingly, it features heavily in this Research Topic. Sexual dimorphism (masculinity in male faces and femininity in female faces) is one of the most studied aspects of facial morphology (for review, see, e.g., Little et al., 2011) and its importance in mate relevant choices is again highlighted in the current issue. Facial morphological correlates of bodily muscle mass in male faces predict female perceptions of facial masculinity; and such correlates are perceived as more attractive for short-term relationships, and less attractive for long-term relationships by said women (Lei et al.). Chen et al. present a meta-analysis of preferences for (primarily) facial sexual dimorphism, confirming well-reported associations between sexual dimorphism and perceived attractiveness. Preferences for highly sexually dimorphic partners are thought to be condition-dependent, reflecting a trade-off made by higher quality individuals, which compromises kind, caring personalities (especially in less masculine men) in favor of more sexually dimorphic, genetically robust individuals. Consistent with this theory, Chen et al. also report a reliable (though small) positive association between own attractiveness and preferences for sexual dimorphism (especially in women rating men as long-term partners, and men rating women as short-term partners). Štěrbová et al. reported that women's long-term partner preferences are generally highly stable from one relationship to the next, with one exception: facial masculinity. The facial masculinity of participants' long-term partners varied more than expected by random coupling—the only trait of 21 measured to do so. In addition, long-term partners with whom the participants had children also had more masculine faces than non-fathers.

OPEN ACCESS

Edited and reviewed by:

Peter Karl Jonason,
University of Padova, Italy

*Correspondence:

Danielle Sulikowski
danielle.sulikowski@gmail.com

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 29 February 2020

Accepted: 17 March 2020

Published: 08 April 2020

Citation:

Sulikowski D, Tan KW, Jones AL,
Welling LLM and Stephen ID (2020)
Editorial: Perceptions of People: Cues
to Underlying Physiology and
Psychology. *Front. Psychol.* 11:643.
doi: 10.3389/fpsyg.2020.00643

Beyond sexual dimorphism, facial adiposity, and skin color are also important cues to physical and physiological health. Facial adiposity is a highly reliable indicator of body mass index and numerous health outcomes (most notably those associated with being overweight or obese; de Jager et al.). Across cultures and ethnicities, skin yellowness (which indicates the presence of antioxidant carotenoids), and to a lesser extent skin redness (an indicator of blood flow associated with physical fitness and levels of sex hormones) are subjectively perceived by observers as indicators of physical health (Tan and Stephen). Cues to physiological health, especially anaerobic capacity, may also be present in human faces. Judgements of the fighting ability of mixed martial arts fighters, based on 360° headshots, were jointly predicted by the fighters' weight and anaerobic capacity (Trebeck et al.). Although this study took advantage of the additional information available in a 360° headshot compared to a single front-on portrait, Trebeck et al. also demonstrated that attractiveness and formidability judgements of front-on, profile, and 360° portraits differ very little in terms of their means, and are strongly inter-correlated. This observation bolsters many conclusions drawn from across the field of evolutionary psychology of face perception, many of which are based on judgements of single front-on portraits.

The judgments we make of human faces may also influence how we perceive non-human primate faces. Primate faces that are least like human faces are perceived as the most beautiful (Rádlová et al.), an effect attributed to the uncanny valley (i.e., the hypothesized relationship between resemblance to humans and emotional response; e.g., Mori et al., 2012). The more like human faces primate faces become, however, the more that human-face predictors of attractiveness judgements (in terms of the exact arrangements of the internal features), also predicted the beauty judgements of the primate faces.

New data-driven methods for analyzing facial morphology have also been presented in this Research Topic. Mogilski and Welling present a novel application of data-driven conjoint analysis showing that eyebrow thickness, jaw prominence, and facial height exhibit superior signaling capacity for judgements of sexual dimorphism and attractiveness compared to other regions of the face. Kleisner et al. proposed a new method for measuring cultural typicality and cross-cultural distinctiveness of faces. Based on calculating an individual face's exact position on the vector connecting the facial morphological averages of the two cultures being compared, Kleisner et al. were able to predict significant variance in faces' perceived distinctiveness. They also demonstrated that in-group/out-group judgements were easier (in a two-alternative forced-choice design) when the out-group face was furthest away from the in-group mean. Both of these novel methodologies open the door for more nuanced analyses of the signaling capacity of human facial morphology.

Non-morphological, behavioral cues, including vocalizations and body odor, also provide information about their bearer. Single men, compared to partnered men, exhibit stronger body odor, and based on their body odor alone are judged by women to be more masculine (Mahmut and Stevenson). Listeners also adjust their perceptions of another based on the pitch of vocalizations. Mixed martial arts fighters are judged to be more

formidable if their roars (spontaneous vocalizations made when victorious) are higher pitched, and also if their speaking voice is lower pitched (Šebesta et al.). Similarly contrasting effects of pitch are observed when comparing singing and speaking voices. The perceived attractiveness of singing and speaking strongly correlate across individuals, and both predict physical size in men (Valentova et al.). Also in men, low-pitched speech, but higher-pitched singing, predict higher sociosexuality. In women, shorter apparent vocal tract length (indicated by shorter spacing between the first four formant frequencies) in speech, but longer apparent vocal tract length during singing, predicted higher sociosexuality (Valentova et al.). It therefore appears that both pitch and range of vocalizations provide important cues as to the quality of potential mates and competitors.

Extended phenotypic traits, manifesting beyond the bearer's physical body, may also constitute socially meaningful cues. Two such adornments include clothing and make-up. Gouda-Vossos et al. investigated how wearing business vs. casual attire influenced perceptions of men's and women's socio-economic status. Business attire increases the perceived economic status of men more than it does for women, while the perceived economic status of women is increased if depicted in business attire alongside a group of men (whereas the economic status of men is not increased by being depicted among a group of women). The consequences of such judgements for interpersonal interactions are not known. Women use cosmetics to alter their physical appearance and this can affect their perceived attractiveness. Batres et al. observed that women using the contraceptive pill spend less time putting on make-up for an outing than do naturally cycling women, and that such women are indeed perceived as wearing more make-up. This finding adds to the breadth of behaviors in which circulating sex hormones are implicated (see Welling and Shackelford, 2019), although the complexities of female cosmetics use, what it signals, and how it is perceived remain to be thoroughly investigated.

The methods used by evolutionary psychologists to understand how the multitude of physical and behavioral cues we possess are signaled and received were also a focus of critique within this special issue (see Bovet; Kleisner et al.; Trebeck et al.). Bovet reviewed the literature concerning women's waist-to-hip ratios and criticized the weak theoretical foundations of the area. Highlighting the large number of potential signaling functions of the waist-to-hip ratio and the small number of studies designed to specifically differentiate between the competing theories, Bovet cautions against rushing into empirical work without first establishing a clear theoretical basis to guide empirical robustness and consistency. Doing so can lead to imprecise, untestable predictions and seemingly contradictory results, leading to the premature rejection of potentially legitimate theories based on findings of objectively low evidentiary value.

Cues and signals of physiology, behavior, and personality exist in faces, bodies, voices and our extended phenotypes. For many such potential cues, we have demonstrated objective links between the putative cue and the underlying quality it may indicate; observed receiver sensitivity to such cues; or put forward coherent adaptive and mechanistic theories accounting

for how such cues may come to signal these qualities in the first place (in either a proximate or ultimate sense). For this area of person perception to continue to move forward, concerted efforts are needed to address all three of the above outcomes for each individual cue-quality-receiver system. By triangulating objective relationships between cues and qualities, the impact of such cues on receiver psychology, and sophisticated mechanistic, functional, and adaptive theories to explain the evolution and

maintenance of these systems, we will amass a body of work with great potential impact for understanding modern human social, romantic, and sexual interactions.

AUTHOR CONTRIBUTIONS

DS was the primary author of the first draft. All authors contributed to revisions and corrections of the final manuscript.

REFERENCES

- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. B* 366:1638. doi: 10.1098/rstb.2010.0404
- Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [From the Field]. *IEEE Robot. Autom. Magaz.* 19:98. doi: 10.1109/MRA.2012.2192811
- Welling, L. L. M., and Shackelford, T. K. (2019). *The Oxford Handbook of Evolutionary Psychology and Behavioral Endocrinology*. Oxford University Press: New York, NY.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sulikowski, Tan, Jones, Welling and Stephen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Evidence That the Hormonal Contraceptive Pill Is Associated With Cosmetic Habits

Carlota Batres^{1,2*}, Aurélie Porcheron^{3,4}, Gwenaél Kaminski^{5,6}, Sandra Courrèges³, Frédérique Morizot³ and Richard Russell¹

¹ Department of Psychology, Gettysburg College, Gettysburg, PA, United States, ² Department of Psychology, Franklin and Marshall College, Lancaster, PA, United States, ³ CHANEL Fragrance & Beauty Research & Innovation, Pantin, France, ⁴ Laboratoire de Psychologie et NeuroCognition, Université Pierre Mendès-France, Grenoble, France, ⁵ CNRS (UMR 5263), Cognition, Langues, Langage, Ergonomie, Université de Toulouse, Toulouse, France, ⁶ Institut Universitaire de France, Paris, France

Hormonal contraception is known to cause subtle but widespread behavioral changes. Here, we investigated whether changes in cosmetic habits are associated with use of the hormonal contraceptive pill. We photographed a sample of women ($N = 36$) who self-reported whether or not they use the contraceptive pill, as well as their cosmetic habits. A separate sample of participants ($N = 143$) rated how much makeup these target women appeared to be wearing. We found that women not using the contraceptive pill (i.e., naturally cycling women) reported spending more time applying cosmetics for an outing than did women who use the contraceptive pill. We also found that the faces of these naturally cycling women were rated as wearing more cosmetics than the faces of the women using the contraceptive pill. Thus, we found clear associations between contraceptive pill use and makeup use. This provides evidence consistent with the possibility that cosmetic habits, and grooming behaviors more generally, are affected by hormonal contraception.

Keywords: cosmetics, makeup, contraception, birth control, grooming behaviors

INTRODUCTION

The hormonal contraceptive pill is used by approximately 100 million women worldwide (Christin-Maitre, 2013). Its main function is to change the hormonal state of the menstrual cycle in order to mimic, and thus prevent, pregnancy (Alvergne and Lummaa, 2010). While the majority of women take the contraceptive pill to prevent pregnancy, approximately 14% of women use it for other reasons, such as for lessening menstrual pain and migraines (Cooper and Adigun, 2017).

In addition to its medical side effects, the contraceptive pill has also been linked to several behavioral effects (Welling et al., 2012). For example, one study found that married women not using the contraceptive pill (i.e., naturally cycling women) showed an increase in female-initiated sexual behavior at the time of ovulation, whereas women using the contraceptive pill did not show such a rise (Adams et al., 1978). Other studies have also found that the hormonal state of the contraceptive pill changes women's partner preferences. For instance, unlike naturally cycling women, one study found that those using the contraceptive pill prefer less masculine male faces (Little et al., 2002, but see Jones et al., 2018).

Research has also found that the contraceptive pill affects women's ability to attract mates. For example, one study found that for naturally cycling women, their voices became more attractive as

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Benedict C. Jones,
University of Aberdeen,
United Kingdom
Danielle Leigh Wagstaff,
Federation University, Australia

*Correspondence:

Carlota Batres
cbatres@fandm.edu

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 May 2018

Accepted: 24 July 2018

Published: 23 August 2018

Citation:

Batres C, Porcheron A, Kaminski G,
Courrèges S, Morizot F and Russell R
(2018) Evidence That the Hormonal
Contraceptive Pill Is Associated With
Cosmetic Habits.
Front. Psychol. 9:1459.
doi: 10.3389/fpsyg.2018.01459

their risk of conception increased but no such effect was found for women using the contraceptive pill (Pipitone and Gallup, 2008). One field study examined the tip earnings of professional lap dancers and found that those who were naturally cycling made more money per shift than those who were using the contraceptive pill (Miller et al., 2007).

In this study, we aimed to examine whether the use of the contraceptive pill also influences grooming behaviors, particularly cosmetic habits. Previous research has found that among naturally cycling women, the time spent putting on cosmetics, and the rated level of cosmetics used by women, seems to be higher near ovulation (Guéguen, 2012). No study, however, has examined whether there is a difference in cosmetic habits between naturally cycling women and women using the contraceptive pill. Previous research has shown that, unlike women using the contraceptive pill, naturally cycling women change their appearance near ovulation to look more attractive (Haselton et al., 2007; Durante et al., 2011). Given that makeup is one way that women can alter their attractiveness (Graham and Jouhar, 1981; Jones et al., 2015; Jones and Kramer, 2016; Batres et al., 2018), we predicted that naturally cycling women would report spending more time applying cosmetics, and would be rated as wearing more cosmetics, than women using the contraceptive pill.

STUDY 1

Methods

Participants and Procedure

Thirty six women (M age = 19.58 years, SD = 2.14) completed the study. The participants were recruited at the University of Grenoble using advertisements and obtained financial compensation to participate in the study. The research was performed in accordance with the Declaration of Helsinki: it was conducted with the understanding and the written consent of each participant (who were instructed that their photographs would be taken) and was approved by local ethics boards (CNRS and the University of Grenoble).

After arriving at the laboratory, 20 min were allowed to pass in order to allow the participants' skin to acclimatize to the indoor temperature. Photographs were then taken of each participant facing forward, under constant camera/lighting conditions, with neutral expressions, and closed mouths. Each participant also completed a questionnaire in which she reported whether she was taking hormonal contraceptives, if so, what type of hormonal contraceptives (e.g., the contraceptive pill, an intrauterine device), whether she had regular menstrual cycles, and whether she was in a relationship. All 36 women reported having regular menstrual cycles. Twenty women reported not using any hormonal contraceptives (i.e., naturally cycling) and 16 women reported using hormonal contraceptives, out of which 100% reported using the contraceptive pill.

Each participant also completed two questions pertaining to her cosmetic habits: how much time she spent on her daily makeup in the morning and how much time she spent making up for an outing. These two questions allowed us to define two

variables concerning cosmetic habits: daily cosmetics and outing cosmetics. To examine the association between contraceptive pill use and cosmetic habits, we performed independent samples t -tests on both cosmetic habits variables.

Results

Naturally cycling women reported spending more time applying cosmetics (both daily and for an outing) than women using the contraceptive pill (see **Figure 1**). The difference between naturally cycling women and women using the contraceptive pill was not statistically significant for the amount of time they spent on their daily cosmetics, $t(33) = 1.59$, $p = 0.121$, Cohen's $d = 0.51$, but it was statistically significant for the amount of time they spent on their cosmetics for an outing, $t(19.6) = 2.30$, $p = 0.033$, Cohen's $d = 0.80$.

STUDY 2

Methods

Participants and Procedure

One hundred and forty three Gettysburg College students participated in Study 2 (M age = 18.53 years, SD = 0.87, 56 male, 87 female) as part of a course requirement. Ethical approval was received from the Gettysburg College Institutional Review Board. Participants were instructed that they would be viewing and rating face images on a computer. Participants were asked to rate each face on the question: "How much makeup does this face have?," where 1 = Very little makeup and 7 = A lot of makeup. The faces presented were those of the women from Study 1. To examine the association between pill use and perceived amount of cosmetics, we performed a linear mixed model.

Results

Due to repeated measurements, we included the target faces and the participants as random effects. Naturally cycling women

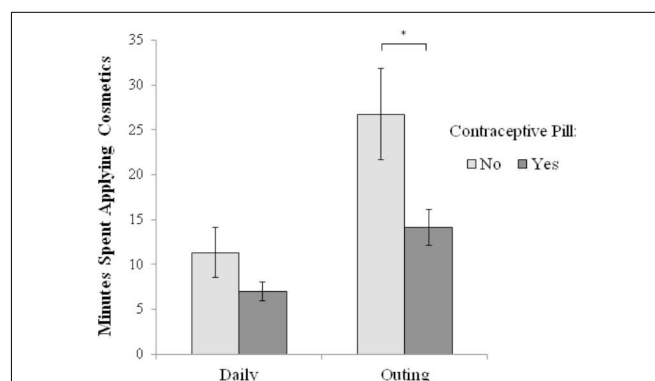


FIGURE 1 | In minutes, reported amount of time spent applying daily cosmetics in the morning and reported amount of time spent making up for an outing, for both naturally cycling women (i.e., those not using the contraceptive pill) and women using the hormonal contraceptive pill. The asterisk indicates a significant difference (* $p < 0.05$). Error bars indicate the standard error of the mean.

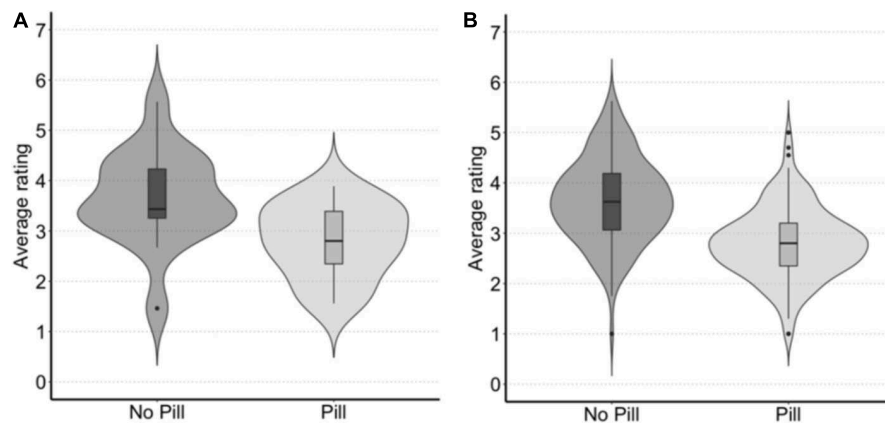


FIGURE 2 | Average rated amount of cosmetics (where 1 = Very little makeup and 7 = A lot of makeup) by (A) target faces and (B) participants, for both naturally cycling women (i.e., those not using the contraceptive pill) and women using the contraceptive pill.

were rated as having higher levels of cosmetics [3.65 (3.24–4.06)] than women using the contraceptive pill [2.79 (2.43–3.17)] (see **Figure 2**). We found that the difference in amount of perceived cosmetics between naturally cycling women and women using the contraceptive pill was statistically significant [$\chi^2(1) = 8.86$, $p = 0.003$, $\beta = 0.85$ [0.31–1.39]].

DISCUSSION

The results from Studies 1 and 2 provide the first evidence that the use of the hormonal contraceptive pill is associated with cosmetic habits. In Study 1, we found that naturally cycling women spent more time applying cosmetics for an outing than women using the contraceptive pill. On average, naturally cycling women reported spending an extra 13 min applying cosmetics for an outing than women using the contraceptive pill. Naturally cycling women also reported spending more time applying their daily cosmetics than women using the contraceptive pill, but this difference was not statistically significant. In Study 2, we found that the faces of the naturally cycling women (taken during the day) were rated by participants as having higher amounts of cosmetics than the faces of the women using the contraceptive pill. This suggests that while naturally cycling women may not spend more time applying their daily makeup, they use more visible cosmetics. These findings thus provide evidence that contraceptive pill use is associated with cosmetics use.

This is consistent with previous research suggesting that naturally cycling women are found more attractive than women using the hormonal contraceptive pill (Miller et al., 2007). Our study, however, proposes that part of such difference in attractiveness may be due to cosmetics. In other words, naturally cycling women may, in part, be found more attractive than women using the contraceptive pill because they wear more cosmetics, which greatly increases attractiveness (Graham and Jouhar, 1981; Batres et al., 2018). For instance, Miller et al.'s (2007) study found that naturally cycling lap dancers had higher tip earnings than those using the contraceptive pill, but

part of that difference may be explained by a change in the amount of cosmetics worn by the dancers. Indeed, research has observed that waitresses receive higher tips when they are wearing cosmetics compared to when they are not (Jacob et al., 2009). This suggests it is likely that different makeup use explains part of the differences in attractiveness found between naturally cycling women and women using hormonal contraception.

Our results may also point to a larger issue: that the contraceptive pill may be associated with behavioral changes that affect women's grooming practices more generally, of which cosmetics is just one part. In other words, the contraceptive pill may suppress how women adorn themselves, otherwise referred to as the extended phenotype (Etcoff et al., 2011). Some studies have examined how the extended phenotype changes throughout the menstrual cycle. For instance, one study found that self-grooming and ornamentation through attractive choice of dress increased during high fertility periods (Haselton et al., 2007). However, a recent longitudinal study did not find evidence for fertility-linked changes in women's clothing choices (Arslan et al., 2017). Similar studies are needed to compare the extended phenotype of women who are using the contraceptive pill and those that are not. Such research would shed light on whether cosmetic habits are the only form of ornamentation influenced by the contraceptive pill or whether it is just one facet of a greater behavioral shift in how women present themselves (e.g., grooming, jewelry, clothing).

Our study did not control the phase of the menstrual cycle in which women were photographed nor asked how long the women had been using or not using the contraceptive pill and therefore future research should control for this. We also did not define what was meant by an outing and the interpretation of this could have had some impact on self-reported time spent applying cosmetics. For example, women may have reported less time if an outing was interpreted as going out for dinner versus if it was interpreted as going out to a club. Future studies would therefore benefit from being more specific when asking about time spent applying cosmetics for an outing. Similarly, while we asked the women how much time they spend on their daily makeup in

the morning, it would have been helpful to ask how much time they spent making up that exact morning in order to better link cosmetics application time with perceptual differences. Moreover, adding other measures of cosmetics use would also be beneficial in future studies (e.g., “How natural does the makeup on this face look?”). Women who are skilled at perfecting the “natural” look are likely to spend more time applying cosmetics but may appear to be wearing less makeup.

Further research is still needed to better understand links between the contraceptive pill and cosmetic habits as our study was a non-randomized between-subjects design and research has found that naturally cycling women and women using the contraceptive pill differ in other ways (Alexander et al., 1990). For example, Little et al. (2002) found that women using the contraceptive pill reported having more lifetime sexual partners than naturally cycling women. A related concern is that in our study only 19% of naturally cycling women reported being in a relationship, while 85% of women using the contraceptive pill reported being in a relationship. Our sample was too small to include relationship status as a factor in our analyses and thus future research with a larger sample is needed. No study has examined differences in cosmetic habits between single women and women in a relationship, however, such a difference could very well be possible. In order to address these two concerns and confirm the link between the contraceptive pill and cosmetic habits, an experimental design, rather than the correlational one used here, would be needed. More specifically, women would need to be randomly assigned to use or not use the contraceptive pill in order to be able to establish a casual effect between the contraceptive pill and cosmetic habits. However, such an experiment is unlikely due to ethical reasons, which is why the current literature has relied on non-randomized between-subjects designs (e.g., Adams et al., 1978; Little et al., 2002; Miller et al., 2007; Pipitone and Gallup, 2008; Welling et al., 2012).

CONCLUSION

In conclusion, we found that, compared to women using the contraceptive pill, naturally cycling women self-reported

spending more time applying cosmetics for an outing, and their faces were rated as having higher levels of cosmetics. These results provide initial evidence that the contraceptive pill is associated with cosmetic habits. Moreover, this association may be part of a broader relationship between contraceptive pill use and other grooming behaviors, stemming from hormone-mediated changes in motivations.

ETHICS STATEMENT

The experiments were undertaken with the understanding and written consent of each subject, with the approval of the appropriate local ethics committees, and in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

AUTHOR CONTRIBUTIONS

CB, AP, GK, FM, and RR conceived and designed the research. CB and GK acquired the data. CB, GK, and RR analyzed the data. CB, AP, GK, SC, and RR interpreted the results. CB drafted the work. CB, AP, GK, SC, and RR critically revised the paper. CB, AP, GK, SC, FM, and RR approved the final version to be published and agreed to be accountable for the content of the work.

FUNDING

The research was funded in part by CHANEL Fragrance & Beauty Research & Innovation.

ACKNOWLEDGMENTS

We thank Lucille Brunero for help with collecting data for this manuscript.

REFERENCES

- Adams, D. B., Gold, A. R., and Burt, A. D. (1978). Rise in female-initiated sexual activity at ovulation and its suppression by oral contraceptives. *N. Engl. J. Med.* 299, 1145–1150. doi: 10.1056/NEJM197811232992101
- Alexander, G. M., Sherwin, B. B., Bancroft, J., and Davidson, D. W. (1990). Testosterone and sexual behavior in oral contraceptive users and nonusers a prospective study. *Horm. Behav.* 24, 388–402. doi: 10.1016/0018-506X(90)90017-R
- Alvergne, A., and Lummaa, V. (2010). Does the contraceptive pill alter mate choice in humans? *Trends Ecol. Evol.* 25, 171–179. doi: 10.1016/j.tree.2009.08.003
- Arslan, R. C., Schilling, K. M., Gerlach, T. M., and Penke, L. (2017). Using 26 thousand diary entries to show ovulatory changes in sexual desire and behaviour. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/jp2ym
- Batres, C., Russell, R., Simpson, J. A., Campbell, L., Hansen, A. M., and Cronk, L. (2018). Evidence that makeup is a false signal of sociosexuality. *Pers. Individ. Dif.* 122, 148–154. doi: 10.1016/j.paid.2017.10.023
- Christin-Maitre, S. (2013). History of oral contraceptive drugs and their use worldwide. *Best Pract. Res. Clin. Endocrinol. Metab.* 27, 3–12. doi: 10.1016/j.beem.2012.11.004
- Cooper, D. B., and Adigun, R. (2017). *Oral Contraceptive Pills. Treasure Island.* Orlando, FL: StatPearls Publishing.
- Durante, K. M., Griskevicius, V., Hill, S. E., Perilloux, C., and Li, N. P. (2011). Ovulation, female competition, and product choice: hormonal influences on consumer behavior. *J. Consum. Res.* 37, 921–934. doi: 10.1086/656575
- Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., and House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: modulation of the perception of biologically important facial signals. *PLoS One* 6:e25656. doi: 10.1371/journal.pone.0025656
- Graham, J. A., and Jouhar, A. J. (1981). The effects of cosmetics on person perception. *Int. J. Cosmet. Sci.* 3, 199–210. doi: 10.1111/j.1467-2494.1981.tb00283.x
- Guéguen, N. (2012). Makeup and menstrual cycle: near ovulation, women use more cosmetics. *Psychol. Rec.* 62, 541–548. doi: 10.1007/BF03395819

- Haselton, M. G., Mortezaie, M., Pillsworth, E. G., Bleske-Rechek, A., and Frederick, D. A. (2007). Ovulatory shifts in human female ornamentation: near ovulation, women dress to impress. *Horm. Behav.* 51, 40–45. doi: 10.1016/j.yhbeh.2006.07.007
- Jacob, C., Guéguen, N., Boulbry, G., and Ardiccioni, R. (2009). Waitresses' facial cosmetics and tipping: a field experiment. *Int. J. Hosp. Manag.* 29, 188–190. doi: 10.1016/j.ijhm.2009.04.003
- Jones, A. L., and Kramer, R. S. (2016). Facial cosmetics and attractiveness: comparing the effect sizes of professionally-applied cosmetics and identity. *PLoS One* 11:e0164218. doi: 10.1371/journal.pone.0164218
- Jones, A. L., Russell, R., and Ward, R. (2015). Cosmetics alter biologically-based factors of beauty: evidence from facial contrast. *Evol. Psychol.* 13, 210–229. doi: 10.1177/147470491501300113
- Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., et al. (2018). No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. *Psychol. Sci.* 29, 996–1005. doi: 10.1177/0956797618760197
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. Biol. Sci.* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Miller, G., Tybur, J. M., and Jordan, B. D. (2007). Ovulatory cycle effects on tip earnings by lap dancers: economic evidence for human estrus? *Evol. Hum. Behav.* 28, 375–381. doi: 10.1016/j.evolhumbehav.2007.06.002
- Pipitone, R. N., and Gallup, J. G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evol. Hum. Behav.* 29, 268–274. doi: 10.1016/j.evolhumbehav.2008.02.001
- Welling, L. L. M., Puts, D. A., Roberts, S. C., Little, A. C., and Burriss, R. P. (2012). Hormonal contraceptive use and mate retention behavior in women and their male partners. *Horm. Behav.* 61, 114–120. doi: 10.1016/j.yhbeh.2011.10.011

Conflict of Interest Statement: AP, FM, and SC work at CHANEL Fragrance & Beauty Research & Innovation, a cosmetics company, and CB and RR receive funding from CHANEL Fragrance & Beauty Research & Innovation.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Batres, Porcheron, Kaminski, Courrèges, Morizot and Russell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Relationship Between Observers' Self-Attractiveness and Preference for Physical Dimorphism: A Meta-Analysis

Lijun Chen^{1,2}, Xiaoliu Jiang¹, Huiyong Fan³, Ying Yang¹ and Zhihong Ren^{4,5*}

¹ School of Humanities and Social Sciences, Fuzhou University, Fuzhou, China, ² Institute of Psychological and Cognitive Sciences, Fuzhou University, Fuzhou, China, ³ College of Teacher, Bohai University, Jinzhou, China, ⁴ Key Laboratory for Adolescent Cyberpsychology and Behavior of the Education Ministry, Wuhan, China, ⁵ Laboratory of Human Development and Mental Health, Institute of Psychology, Central China Normal University, Wuhan, China

OPEN ACCESS

Edited by:

Lisa L. M. Welling,
Oakland University, United States

Reviewed by:

Anthony Little,
University of Bath, United Kingdom
Nicole Barbaro,
Oakland University, United States
Lei Chang,
University of Macau, China

*Correspondence:

Zhihong Ren
psyren@qq.com

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 15 June 2018

Accepted: 19 November 2018

Published: 05 December 2018

Citation:

Chen L, Jiang X, Fan H, Yang Y and
Ren Z (2018) The Relationship
Between Observers'
Self-Attractiveness and Preference for
Physical Dimorphism: A
Meta-Analysis.
Front. Psychol. 9:2431.
doi: 10.3389/fpsyg.2018.02431

Background: Many studies have reported an association between observers' self-attractiveness and their preference for sexual dimorphism across different physical domains, including the face, voice, and body. However, the results of these studies are inconsistent. Here, a meta-analysis was conducted to estimate the association between observers' own attractiveness and their dimorphic preference.

Methods: Major electronic databases including PsycINFO, Web of Science, PubMed, ProQuest, and Google Scholar were searched during April 2017 (the first time) and April 2018 (the second time). The effect size computation and moderating effect analyses were conducted separately for masculine and feminine preferences.

Results: We identified 5,359 references, of which we included 25 studies ($x = 55$, x = number of the effect size) with 6,853 participants in the meta-analysis. Across these studies, the correlation between observers' own attractiveness and their sexual dimorphic preference was 0.095 ($x = 55$) and that for preference for masculinity ($x = 39$) and femininity ($x = 16$) were 0.102 and 0.076, respectively. The results of the funnel plot, Egger's regression method, and fail-safe number suggested that there was no obvious publication bias. The relationship depended on the relationship context (short or long-term), opposite or same sex (the gender of the observer and host), measures of observers' self-attractiveness (subject or objective), and preference task (e.g., attractiveness rating, forced-choice, and face sequence test). Furthermore, for female participants, using a hormonal contraceptive also influenced their masculinity preference. The effect size for the preference for a masculine body and voice was larger than that for facial masculinity.

Conclusion: We found a small but significant correlation between self-attractiveness and physical dimorphic preference, the relationship was moderated by the relationship context, same/opposite-sex, and contraceptive using. These three moderating effects represented the observer's trade-off on good genes, good provider and good father (3Gs) consistent with the life history strategies. Besides, measurement of observers' attractiveness, type of preference task and stimuli may also involve the relationship.

Keywords: femininity, masculinity, meta-analysis, self-attractiveness, sexual dimorphism

INTRODUCTION

Secondary sexual characteristics in adult humans reflect the masculinization or feminization that occurs during puberty (Perrett et al., 1998; Rhodes, 2006). Physical sexual dimorphism is a broad concept that could include sexual dimorphism in multiple domains (e.g., face, body, voice). Sexually dimorphic physical traits are important for mate choice and mate preference in many species, including humans. Several previous studies have observed that humans' preferences for physical cues of extreme secondary sexual characteristics (more feminine for women, more masculine for men) in different domains (e.g., visual, vocal, and bodily) are correlated (Little et al., 2007; Fraccaro et al., 2010). These correlations demonstrate a systematic, rather than arbitrary, variation in humans' preferences for sexual dimorphism, which are consistent with the proposal that sexually dimorphic cues in different domains reflect a common underlying aspect of quality. On the evolutionary view, femininity of women, masculinity of men are proposed to be more attractive because they advertise the good genes of an individual (Rhodes, 2006). Among humans, physical characteristics consistent with the owner's gender are correlated with indices of long-term health (Rhodes et al., 2003; Thornhill and Gangestad, 2006), reproductive potential (Puts, 2005; Rhodes et al., 2005), and low parasite loadings and high immune competence (Thornhill and Gangestad, 1993, 1996), but negatively correlated with prosociality (Haselton, 2005; Haselton and Gangestad, 2006). Men's masculine traits indicate untrustworthiness and bad parental traits (Boothroyd et al., 2007; Smith et al., 2009b), and women's femininity are considered as more likely to be unfaithful, to pursue short-term relationships, and to be in higher risk of cuckoldry (Little et al., 2014).

According to the life history (LH, referring to organisms capturing energy from the environment and using it to produce more organisms) trade-off model strategies in mating choice (Gangestad and Simpson, 2000; Del Giudice and Belsky, 2011), women focused on two types of characteristics when they chose a mate: those indicating a "good provider" (social-economic characteristics, such as wealth, education, career) and "good genes" (physical characteristics) (Gangestad and Buss, 1993; Gangestad and Simpson, 2000). Other researchers believed that the framework of women's mate preference should involve three Gs. Besides good genes, good providers, women also prioritize a man's personality traits—for example, being kind, loving, and staying at home—that constitute good fathers (Buss and Shackelford, 2008; Lu et al., 2015). Good provider indicate men have resource to invest in parenting, traits of good father can reflect men's intention to help raising young, both of the two types of characteristics represent parenting of reproductive effort in post mating events (parental investment), otherwise good genes characteristics were realized as mating attributes affecting premating decisions. When exposed to contrasting environments, women would have evolved to make trade-offs between investment qualities and indicators of good genes contingent on specific environmental conditions, because good genetic males tend to have more mates at the same time, and they invest in each female than in males of lower phenotypic

quality, women's emphasis on good genes may be at the cost of men's parental investment (Gangestad and Simpson, 2000). Men also encounter trade-off problems, because, while feminine females possess high attractiveness and good genes, they are also associated with negative personality characteristics (unfaithful) (Haselton, 2005; Haselton and Gangestad, 2006), and men also have expectations about maternal investment. For both men and women (but more for men than women), parental warmth and care (good father or mother) correlated negatively with good-gene and good-provider mate values (Chang et al., 2017). Both men and women encounter the tradeoff of 3Gs mating framework.

Therefore, preferences for dimorphism represent the result of trading off between good genes and parental investment. Differences in how humans resolve this trade-off can lead to individual differences in sexual dimorphic preference. For example, attractive women demonstrate stronger preferences for masculine men than relatively unattractive women do (Little et al., 2001; Little and Mannion, 2006). These are called condition-dependent preferences. In the evolution of species, condition-dependent preferences have been observed in many species, in which individuals in good physical condition tend to show stronger preferences for high-quality mates (Bakker et al., 1999). Condition-dependent preferences in both humans and non-humans may have a common function and they may occur because individuals in good physical condition (i.e., attractive individuals) are better able to compete for and/or retain high-quality mates (Little et al., 2001). Additionally, they can offset the costs of choosing a partner with good genes (e.g., by being able to replace a partner more quickly), and therefore, they improve their criteria for mate selection. On the contrary, in order to meet the needs of parental investment, individuals with poor self-conditions may reduce their standards for mate selection, and they prefer mates who are more likely to make high parental investment. The following findings appear to be somewhat analogous to condition-dependent preferences observed in humans. Women's ratings of their own physical attractiveness positively correlated with the strength of their preferences for masculine characteristics in men's faces (Little et al., 2001; Penton-Voak et al., 2003; Smith et al., 2009b). Similar correlations have been found in men's voices (Vukovic et al., 2008, 2010) and bodies (Little et al., 2007). Further, the concept of condition-dependent preferences is conceptualized as "market value dependent preferences": when exposure to attractive same-sex images, women perceived themselves less attractiveness and lower prefer for male facial masculinity; whereas exposure to unattractive same-sex images, they perceived themselves more attractiveness and lower prefer for masculinity (Little and Mannion, 2006).

There is inconsistent evidence on the relationship between observers' own attractiveness and their sexual dimorphic preference. While some experimental results have confirmed that attractive females prefer masculine faces (Smith et al., 2009b; Welling et al., 2009; Kandrik and DeBruine, 2012), others did not find such a relationship (Zietsch et al., 2015; Carrito et al., 2016). However, several studies have found that observers' own attractiveness can interact with other variables and impact their

preferences (Little et al., 2001; Smith et al., 2009a; Burriss et al., 2011; Chen et al., 2017).

Influence of Relationship Context in the Preference of Opposite-Sex

The context in which judgments are made can also contribute to differences in the relationship between dimorphic preference and the observer's self-attractiveness (Little et al., 2001, 2002; Penton-Voak et al., 2003; Carrito et al., 2016). There are two types of relationship context: short-term and long-term. The former refers to a sexual relationship, such as a one-night stand; the latter is a lasting relationship, such as being married. The trade-off theory proposed that contextual factors affect the strength of people's preferences for a masculine or feminine partner (Gangestad and Simpson, 2000). Attractive women prefer more masculine male faces than less attractive women do, and this difference is seen in the context of a long- but not a short-term relationship (Penton-Voak et al., 2003). This result has been supported by several studies across voice and body stimuli (Little et al., 2007; Feinberg et al., 2012). Otherwise, attractive men exhibit stronger preferences for feminine women only in the short-term context (Burriss et al., 2011; Little et al., 2014). As feminine women were seen as more likely to be unfaithful and more likely to pursue short-term relationships (Boothroyd et al., 2008), and the risk of cuckoldry limit men's preferences for femininity in women and that it could additionally lead to preferences for femininity in short-term mates (Little et al., 2014). Thus, we supposed that the influence of attractiveness on male masculine preferences is more pronounced in the long-term than in the short-term context for women, but for men, the effect of attractiveness on female feminine preference is more prominent in the short-term context.

Observer's Age and Contraceptive Using

It has been found that reproductively active women had the strongest preference for males' masculinity (Little et al., 2010; Jones et al., 2011), and another study only observed the age effect in women using no contraception (Little et al., 2002). Some studies have found that hormonal contraceptive use may modulate individual differences in women's masculinity preference (Little et al., 2002; Feinberg et al., 2008; Smith et al., 2009a). Furthermore, Vukovic et al. (2008) found that self-rated attractiveness was positively related to the strength of women's preference for masculinized men's voices in women reporting no use of hormonal contraceptives, but not in those using the same. Actually, effects of female observer's age and contraceptive using both due to their physiological hormone levels. Because females of different ages are in different reproductive stages with different hormone levels. And women using hormonal contraceptives are in a hormonal state similar to pregnancy, and consequently, they are unable to realize the benefits that are thought to be associated with choosing a masculine mate (i.e., increased offspring health) (Smith et al., 2009a). In light of the previous studies, we speculate that attractive-contingent preference is stronger in women who do not use hormonal contraceptives than in those who do.

Measures of Observer's Own Attractiveness

Initially, researchers used self-rated items to determine observers' own attractiveness. Little et al. (2001) used self-reported attractiveness as an indicator of observers' own attractiveness and found that attractive females preferred masculine male faces; Little and Mannion (2006) found that women's subjective impressions of their own market value (i.e., their self-rated attractiveness) is particularly important with reference to the effects of attractiveness on women's masculinity preferences. However, other researchers failed to confirm these findings (Cornwell et al., 2006). Therefore, some researchers have questioned the veracity of subjective assessments, and began to adopt objective measures of self-attractiveness in experiments (Penton-Voak et al., 2003). These objective measures included other-rated facial attractiveness, waist-to-hip ratio (WHR), and body mass index (BMI) (Penton-Voak et al., 2003; Smith et al., 2009b). found that women with high WHR (an indicator of unattractiveness) and/or relatively low other-rated facial attractiveness preferred more feminine male faces when choosing males for a long-term relationship. This result has been confirmed by another study (O'Connor et al., 2012). Then, what kind of measurement (subjective or objective) would be more sensitive to self-attractiveness? In the current study, we examined the moderating role of the measurement of observers' own attractiveness (subjective/objective) and compared the effect sizes of the measurements in this relationship.

Thus, evidence on the relationship between observers' self-attractiveness and preference for sexual dimorphism is equivocal. This is potentially because the preference is condition-dependent. Therefore, we can conclude that observers' self-attractiveness is an important variable that interacts with other variables such as measuring methods, relationship context, same/opposite-sex, and contraceptive use (yes or no) to influence observers' dimorphic preferences. Based on the LH trade-off model strategies and condition-dependent preference, the meta-analysis technique was used in the present meta-analysis to investigate whether observers' sexual dimorphic preferences vary across their self-attractiveness, and to interpret the possible reasons for the divergence. We focused on the following two core issues: What is the totally coefficient of the relationship between the two variables? Which factors moderate their relationship significantly?

METHODS

Information Sources and Search

The following search terms were used in combination: *sexual dimorphism*, *masculin**, *feminin**, *fac**, *bod**, *vocal*, *voice*, and *attractiveness*. Search terms for observers' own attractiveness included *self-rated attractiveness*, *self-perceived attractiveness*, *self-reported attractiveness*, *self-perceptions of attractiveness*, *self-ratings of attractiveness*, *other-rated attractiveness*, and *third-party attractiveness ratings*. Major electronic databases, including PsycINFO, Web of Science, PubMed, ProQuest, and Google Scholar were searched during April 2017(the first time) and April

2018 (the second time). The reference lists of the included studies were searched to identify additional studies.

Eligibility Criteria and Study Selection

Only studies that met the following three criteria were included:

① The relationship between the sexual dimorphic preference and observer's self-attractiveness was investigated in the study. ② Specific data on this relationship were accurately reported in the study (such as the correlation coefficients r ; mean; standard deviation; sample size; or corresponding F , t , or χ^2) to enable the calculation of the effect size, excluding the data of the structural equation model, path analysis, and multivariate regression analysis. In order to avoid missing important literature, we wrote to the authors (first or corresponding author) to obtain the correlation coefficient if it was not reported in the article. ③ In cases where there were multiple reports of the same study, we used the first published report. Two authors independently screened the titles and abstracts of the identified articles to exclude ineligible studies. Disagreements were resolved by discussion. We retrieved the full text of the potentially eligible studies and examined full-text reports for further evaluation. The PRISMA flow diagram (Moher et al., 2010) represents all the steps of the literature search (see **Figure 1**).

Summary Measures

In psychological research, the standard mean difference (d) and correlation coefficient (r) are frequently used to compute effect sizes. In order to integrate the relationship between the sexual dimorphic preference and observers' own attractiveness, the correlation coefficient (r) was used in the present meta-analysis. In some primary studies correlation coefficients can be retrieved from t , F , or χ^2 which reported. The following formulas were used in this context (Card, 2012):

$$r = \sqrt{\frac{t^2}{t^2 + df}}, df = n_1 + n_2 - 2; r = \sqrt{\frac{F(1, -)}{F(1, -) + df(error)}};$$

$$r = \sqrt{\frac{\chi^2}{\chi^2 + N}}.$$

After extracted, all correlations were transformed using Fisher's Z -transformation (Lipsey and Wilson, 2001). The sample distribution of Z_r is approximately equal to the normal distribution (Hittner and Swickert, 2006; Borenstein et al., 2009). The formula for the transformation is as follows: $Z_r = 0.5 \times \ln(\frac{1+r}{1-r})$. The overall Z_r can be computed through weighted Z_r . Then the overall r can be found through an inverse operation of Fisher's Z -transformation (Lipsey and Wilson, 2001; Borenstein et al., 2009). These computation related to overall effect size estimation were conducted under the random-effects model (Borenstein et al., 2009). The specialized statistical software Comprehensive Meta-Analysis (CMA, Version 2.2) was used in the current meta-analysis for conducting all the needed computations and analyses.

Data Extraction

Two authors extracted the following information: first author's name, year of publication, sample size of observers, observers'

age, gender of the stimuli (male or female), observers' gender (male or female), facial attractiveness task (the forced-choice test with an attractiveness rating, the forced-choice test alone, or the face sequence test), sexual dimorphic preference (masculinity or femininity), measures of observers' own attractiveness (objective or subjective), type of stimuli (face, voice, body), and use of contraceptive (yes or no; for female observers only).

Considering that the sexual dimorphic preference is divided into masculine preference and feminine preference, and that the correlation between observers' own attractiveness and their masculine or feminine preferences is opposite, this meta-analysis calculated the effect sizes and conducted moderating effect analyses for masculine and feminine preferences separately. The data is available in **Supplementary Table 1**.

Heterogeneity Test

The heterogeneity test was conducted to test whether the average effect size was heterogeneous. Each effect size of each observation value in this meta-analysis contained real and residual effect sizes, which resulted in the partial false phenomenon of effect sizes. The heterogeneity test of effect sizes is always examined by calculating the Q statistic (Borenstein et al., 2009) and I^2 (Card, 2012). In this systemic review, the two statistical values of I^2 and Q were used to detect the heterogeneity of the included effect sizes.

Publication Bias

Publication bias is a concern for any meta-analytic review because it can lead to a larger combined effect than what actually exists. This type of bias refers to the phenomenon where published studies are more likely to report larger effects. Because studies that have not been published due to their negative or null findings are more difficult to retrieve, and therefore, are less likely to be included in a meta-analysis, an upward bias in the combined effect may occur. Furthermore, English-language publication are more likely to be searched, which leads to an oversampling of statistically significant studies. In this meta-analysis, we use a variety of methods to minimize and test publication bias. When searching the literature, we also searched the most popular and diverse Chinese database CNKI (China National Knowledge Infrastructure), the conference and the dissertation database; and wrote to the important researchers in this field to ask if they had any unpublished research reports. When analyzing the data, we used the funnel plot, Egger's regression and Rosenthal's Fail-safe N to evaluate the publication bias and the degree of its impact (Borenstein et al., 2009).

RESULTS

Description of Studies and Overall Association

Due to the topic involved self-attractiveness and dimorphic preference two variables, we combined 7 terms describing preference (face, voice and body) and 6 terms describing observer's attractiveness to search, our literature searches initially identified 5,359 potential articles from databases, but most of them were unrelated articles; Additionally, in many related studies, they involved the relationship between dimorphic

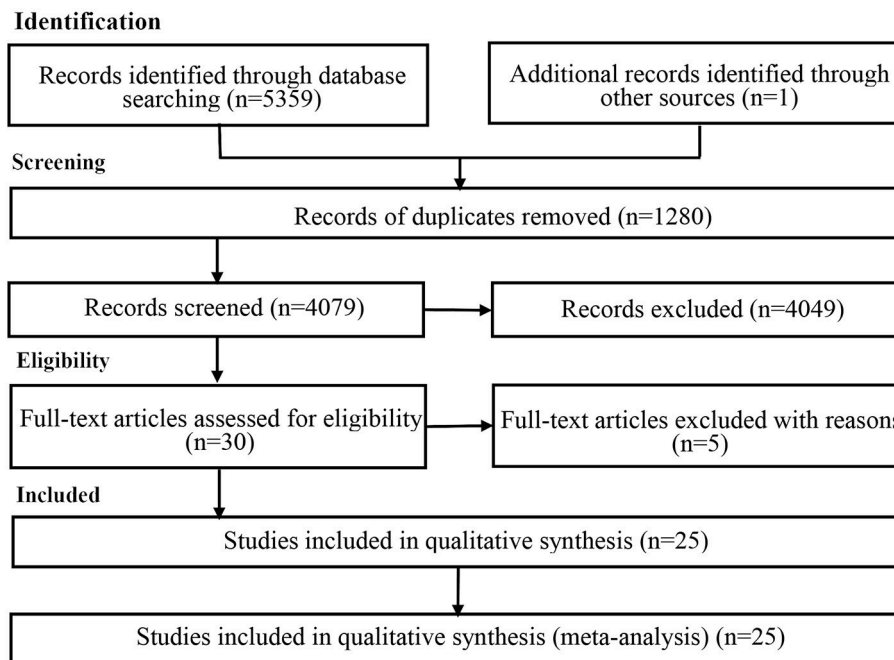


FIGURE 1 | Flow diagram depicting the search protocol and workflow in determining the studies for inclusion in the meta-analysis.

preference and self-attractiveness, whereas, they did not look the relationship as an important topic and not describe in the titles, abstracts, keywords, we also searched the references and citing articles of the related studies, which resulted in many duplicates. Finally the current meta-analysis included 25 studies with 6,853 participants (see **Table 1**). The flow chart has been presented in **Figure 1**. The 25 eligible studies produced 55 effect sizes because 12 studies consisted of multiple datasets. We examined the relationship between preference for sexual dimorphism and observers' own attractiveness. The results showed that the correlation coefficient r ($x = 55$) of the relationship between these two variables was 0.095 (95% CI: 0.059, 0.130; $Z = 5.173$, $p < 0.001$). The correlation coefficient r of the relationship between observers' own attractiveness and preference for masculinity ($x = 39$) and femininity ($x = 16$) were 0.103 (95% CI: 0.060, 0.146; $Z = 4.691$, $p < 0.001$) and 0.076 (95% CI: 0.007, 0.145, $Z = 2.162$, $p = 0.031 < 0.05$), respectively. According to Lipsey and Wilson (2001), an effect size r lower than 0.10 indicates a weak correlation. In order to check the stability of the results of the mean effect size analyses, we conducted a sensitivity analysis, which showed that this meta-analysis did not need to eliminate any of the data that had been included. The data is available in **Supplementary Table 2**.

Heterogeneity Test

The overall heterogeneity test ($x = 55$) showed $Q = 145.567$ ($p < 0.001$), $I^2 = 62.904$, that mean there existed moderate heterogeneity (Higgins et al., 2003). For masculine preferences, the result of the heterogeneity test ($x = 39$) showed $Q = 72.93$ ($p < 0.001$), $I^2 = 47.90$ (moderate heterogeneity). For feminine

preferences, the result of the heterogeneity test ($x = 16$) showed $Q = 71.77$ ($p < 0.001$), $I^2 = 79.10$ (high heterogeneity).

Considering that previous studies have shown heterogeneity of sexual dimorphic preferences (Wood et al., 2014), and according to Borenstein et al. (2009), if the true effect varies across studies using different samples, it is more reasonable to use the random model. A large number of studies have shown that sexual dimorphic preferences are influenced by observers' own attractiveness, age, gender, and sexual orientation (Zheng and Zheng, 2016). Therefore, the random model was more suitable for the present meta-analysis.

Publication Bias

We could not find any Chinese published researches in the CNKI database, and none of the important researchers said they have unpublished researches on this topic. We used the funnel plot, Egger's regression and Rosenthal's Fail-safe N to evaluate the publication bias of the studies included in this meta-analysis. In the absence of publication bias, the studies will be distributed symmetrically about the mean effect size, since the sampling error is random. Otherwise, if the funnel plot is asymmetrical at the bottom, there may be publication bias. The funnel plots in the current analysis are a little asymmetrical at the bottom (see **Figures 2–4**). Because the interpretation of a funnel plot is largely subjective, the Egger's method has been proposed to quantify or test the publication bias (Egger et al., 1997). So, Egger's method was conducted to detect the publication bias, the result is that $t_{(47)} = 1.39$, $p = 0.17$, which means there is no significant bias. Our current meta-analysis reported a significant p -value based on 25 studies. According to Rosenthal's suggestion,

TABLE 1 | Characteristics of the 25 studies included in the present meta-analysis.

References	Task of preference	Sexual dimorphism preference	Measures of observer's own attractiveness	Sample size of observers (n)	Observer's age	Observer's gender	Impact factor
Burriss et al., 2011	Attractiveness rating	Face masculinity	Subjective: self-rated attractiveness	111	20.73 ± 3.37	Men	1.877
	Attractiveness rating	Face masculinity	Objective: other-rated attractiveness	68	–	Men	
Carrito et al., 2016	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	48	22.65 ± 6.60	Women	3.223
	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	61	20.11 ± 4.26	Women	
	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	26	20.39 ± 2.95	Women	
Cornwell et al., 2006	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	46	19.5 ± 1.36	Women	4.997
Feinberg et al., 2012	Sequence test	Voice masculinity	Subjective: self-rated attractiveness	43	–	Women	2.752
Fraccaro et al., 2010	Forced-choice test	Face femininity	Subjective: self-rated attractiveness	178	27.61 ± 9.81	Men	1.278
	Forced-choice test	Voice femininity	Subjective: self-rated attractiveness	178	27.61 ± 9.81	Men	
Holzleitner and Perrett, 2017	Attractiveness rating	Face masculinity	Subjective: self-rated attractiveness	156	18–45	Women	-
Jones et al., 2007	Forced-choice test	Face femininity and masculinity	Subjective: self-rated attractiveness	140	19.84 ± 2.43	Men	2.529
Jones et al., 2010	Forced-choice test	Face masculinity	Subjective: self-rated attractiveness	104	23.3 ± 4.90	Women	1.82
Jones et al., 2011	Forced-choice test	Face femininity	Subjective: self-rated attractiveness	155	48.6 ± 8.1	Women	3.348
Kandrik and DeBruine, 2012	Forced-choice test	Face masculinity	Subjective: self-rated attractiveness	1000	19.5 ± 1.79	Women	1.704
Lefevre and Saxton, 2017	Forced-choice test	Face masculinity	Subjective: self-rated attractiveness	125	20.57 (18–29)	Women	3.223
Little et al., 2001	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	66	22 ± 5.20	Women	3.192
	Face sequence test	Face masculinity	Subjective: self-rated attractiveness	115	22.4±5.4	Women	
Little and Mannion, 2006	Forced-choice test and attractiveness rating	Face masculinity	Subjective: self-rated attractiveness	65	23.5 ± 5.6	Women	2.711
	Forced-choice test and attractiveness rating	Face masculinity	Subjective: self-rated attractiveness	75	21.4 ± 3.4	Men	2.254
Little et al., 2014	Face sequence test	Face femininity	Subjective: self-rated attractiveness	393	27.6 ± 6.5	Men	
Little et al., 2007	Forced-choice test	Body masculinity	Subjective: self-rated attractiveness	97	24.9 ± 5.5	Women	3.401
Moore et al., 2011	Face sequence test	Face femininity	Subjective: self-rated attractiveness	212	23.51 ± 9.54	Women men	1.07
O'Connor et al., 2012	Attractiveness rating	Face masculinity	Objective: other-rated attractiveness	63	18.71 ± 1.71	Women	1.947
Penton-Voak et al., 2003	Face sequence test	Face masculinity	Subjective: self-rated attractiveness Objective: other-rated attractiveness, WHR	82	20.2	Women	1.28
Smith et al., 2009b	Forced-choice test and attractiveness rating	Face femininity	Subjective: self-rated attractiveness	147	19.92 ± 3.55	Women	1.278
Smith et al., 2009a	Forced-choice test	Face masculinity	Objective: WHR and BMI	32	–	Women	1.878
Vukovic et al., 2008	Forced-choice test and	Voice masculinity	Subjective: self-rated attractiveness	58+65	19.72 ± 2.85	Women	1.598
Vukovic et al., 2010	Attractiveness rating	Voice masculinity	Objective:	54+58	19.92 ± 2.36	Women	2.926

(Continued)

TABLE 1 | Continued

References	Task of preference	Sexual dimorphism preference	Measures of observer's own attractiveness	Sample size of observers (n)	Observer's age	Observer's gender	Impact factor
Welling et al., 2008	Forced-choice test and attractiveness rating	Face femininity	Subjective: self-rated attractiveness	43	19.01 ± 1.52	Women	1.598
Welling et al., 2009	Forced-choice test	Face masculinity	Subjective: self-rated attractiveness	808	18.22 ± 1.09	Women	1.878
Zietsch et al., 2015	Forced-choice test	Face masculinity	Subjective: self-rated attractiveness	2160	33.11 ± 5.00	Women	5.476

we should compute how many missing studies we would need to retrieve and incorporate in the analysis before the p -value became non-significant, the number of the missing studies was called Rosenthal Fail-safe N . The larger number of studies that are needed to nullify the effect, the more confident we can be of a real effect (Rosenthal, 1979). The result of Fail-safe N showed that at least 973 studies with the opposite conclusions would be required to overturn the findings of this meta-analysis.

Subgroup Analysis

Given the moderate heterogeneity of effect sizes and influence factors aforementioned in the introduction, we conducted subgroup analyses to examine whether the effect sizes varied according to measures of observers' self-attractiveness, type of stimulus, tasks of preference and same/opposite-sex stimuli. The present meta-analysis included only one effect size of same-sex feminine preferences in male stimuli and none of masculine preferences in female stimuli. Thus, we conducted subgroup analysis on feminine preferences for female stimuli and masculine preferences for male stimuli for the same-opposite sex analysis. Moreover, in masculine preference, all the effect sizes on relationship context and contraceptive use were extracted from the preference of opposite-sex (female preference for male stimulus); in feminine preference, all the data on relationship context also came from the preference of opposite-sex (male preference for female stimulus). The results of the subgroup analysis have been presented in **Table 2**. The moderated analysis showed that, for masculine ($x = 39$), task of preference, measures of observers' self-attractiveness, type of stimulus affected the association between observers' self-attractiveness and their masculine preferences, additionally, women's preference for male masculinity varied across contraceptive use/no and short-long term relationship context. And for feminine preference ($x = 16$), this association also depended on the task of preference, relationship context, same/opposite-sex stimuli.

Meta-Regression Analysis

To assess the influence of observers' age, published year, and sample size of observers on the effect size (r coefficient), considering these three variables are continuous, meta-regression analysis were carried out for masculine preference and feminine preferences separately. Firstly, we looked the effect size of relationship between self-attractiveness and masculine preference as dependent variable. The results showed the effect

size decreased with observers' age, published year, and sample size of observers (see **Table 3**). Specifically, the older observers were, the more negative relationship between self-attractiveness and masculinity preference. In other words, in younger people, individuals in high physical condition would more prefer masculinity. Further, the larger the sample size of the study, the weaker was this relationship, and the later the publishing year of the study, the smaller was the effect. Subsequently, the effect size of relationship between self-attractiveness and feminine preference was used as dependent variable. Only the publishing year significantly positively influenced the relationship.

DISCUSSION

The Overall Association

Human sexual dimorphic preference is condition-dependent, and it is an attractive-contingent preference. However, the empirical evidence is inconsistent. The present meta-analysis provides a quantitative synthesis of the available evidences on the attractive-contingent preference, and reveals a significant overall relationship ($r = 0.095$, $x = 55$). Additionally, self-attractiveness is significantly positively but weakly related with masculine and feminine preference (masculinity: $r = 0.102$, $x = 39$; femininity, $r = 0.076$, $x = 16$). These findings are consistent with the concept of condition-dependent preference. Condition dependence lies at the heart of the trade-off between costly sexual traits and other major fitness components such as survival and growth. Variability among individuals' physical condition can potentially influence the form, direction, and intensity of sexual selection in the population as a whole (Widemo and Sæther, 1999). Some previous studies have confirmed that high-quality females were more attracted to markers of quality in males (masculine men) across different domains (face, voice, body, and smell) (Little et al., 2011b), which is due to the fact that their own high attractiveness means that lower parental investment is less detrimental. Actually, there is a more complex relationship between self-condition and mating choice, as self-attractiveness does not occur in isolation (Little et al., 2014). The objective or subjective measurement of self-attractiveness (Smith et al., 2009b), relationship context of preference for opposite-sex stimuli (Little et al., 2007; Kandrik and DeBruine, 2012), type of stimuli would also play an important role in observers' sexual dimorphic preferences.

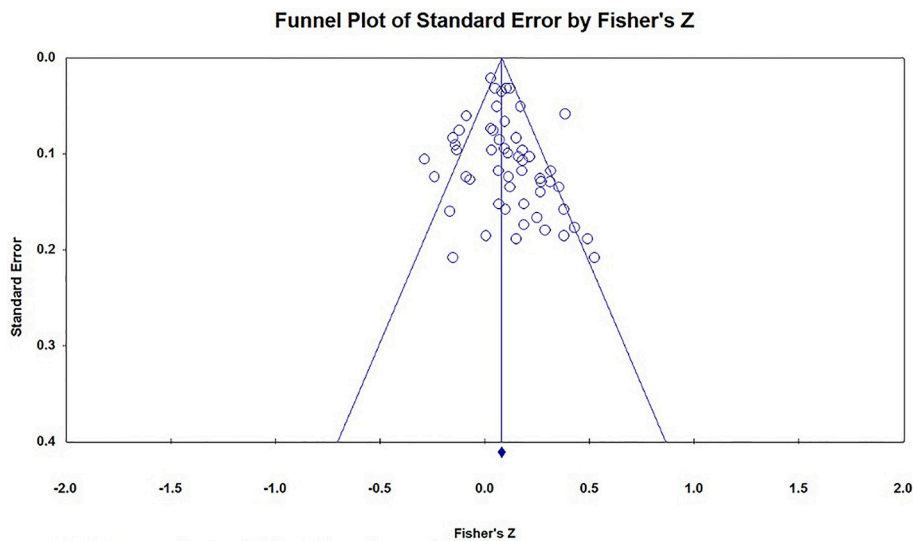


FIGURE 2 | Funnel plot of all the effect size ($n = 55$).

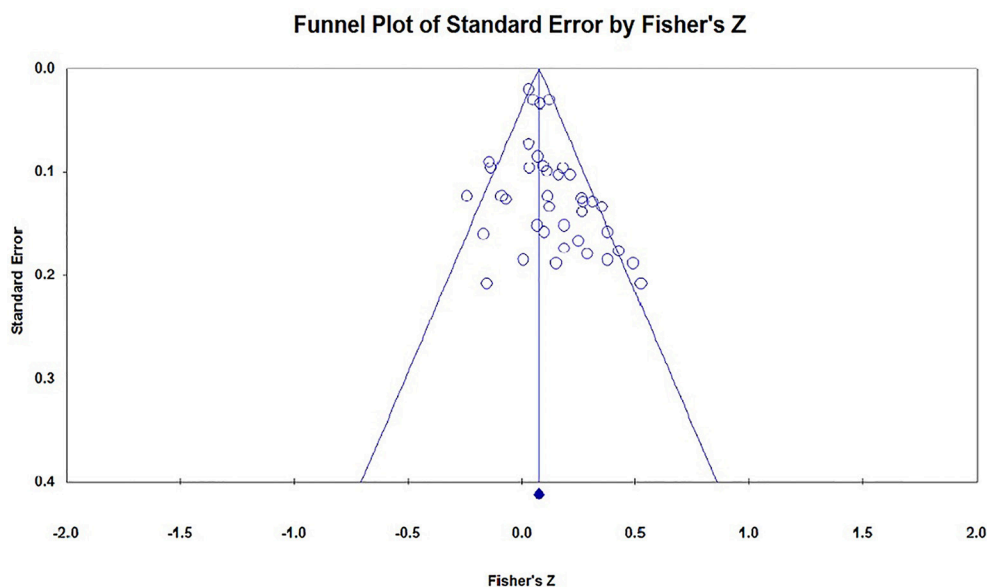
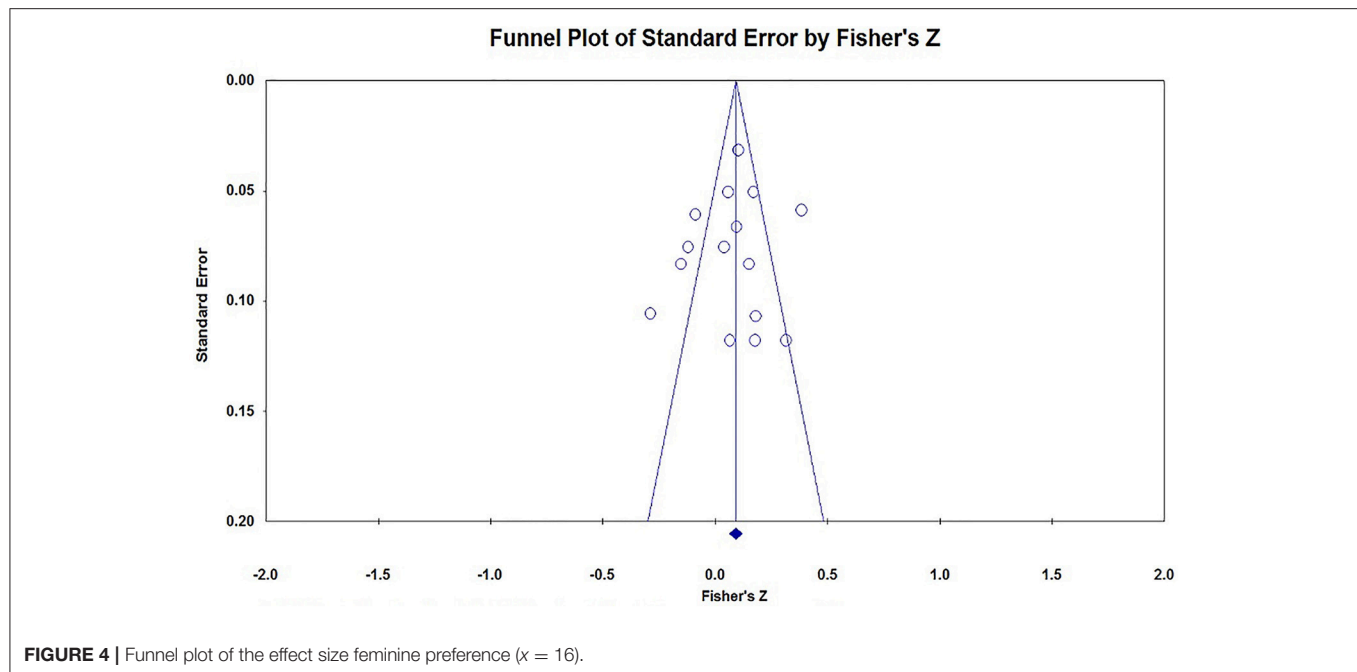


FIGURE 3 | Funnel plot of the effect size of masculine preference ($n = 39$).

The Influence of Same/Opposite-Sex, Relationship Context, Contraceptive Use: Basing on the Life History Tradeoff Strategies

As shown in **Table 2**, the intensity of attractive-contingent preference did change across same/opposite-sex conditions, which is inconsistent with our hypothesis. In masculine preference for male stimulus, the correlation between women's self-attraction and male masculine preference was stronger than that of men's self-attraction (opposite = 0.107, $n = 35$; same = 0.066, $n = 3$). Patterns of women preferences

for men's masculine features are more complex (Burke and Sulikowski, 2010; Holzleitner and Perrett, 2017). On the viewing of LH strategies, good genes, and good provisioning male mate attributes evolved mainly from polygyny: muscularity in men indicating higher immune competence, physical attractiveness, and dominance, competition, high status, but they attract and tend to have more female mates at the same time, and they invest less in each female. In spite of this, attractive women are more confident and believe that they can offset the costs of choosing good genes and good providing partner (Little et al., 2001). Although, attractive men more likely to prefer masculine men as social allies, self-rated sex typicality is the



stronger trait to predict preference for sex-typical physical cue in same-sex faces (Kandrik and DeBruine, 2012), and appears to be more important influence factor of men's preference for same-sex physical cues. Moreover, the market value (good physical condition) is potentially a more important resource for women's mating choice than for men's cooperative choice (Kandrik, 2017). Otherwise, When it comes to the preference for female femininity, the influence of same-sex is greater than that of opposite-sex (same = 0.219, $x = 3$; opposite = 0.083, $x = 7$), in another word, attractive women showed a more obvious trend than men did. Possibly because more attractive women may perceive less threat from other feminine women and more likely to look them as social allies (Fisher, 2004). While Attractive men tend to choose charming women but also trade off the benefit and the cost of cuckoldry. Therefore, men's choices also depend on the context of the relationship (long or short term) (Burriss et al., 2011; Little et al., 2014).

In line with our expectations, on women's preferences in male masculinity, long-term context had a larger effect size as compared to the short-term context (correlation in the short-term context was 0.001, $x = 9$; that in the long-term context was 0.228, $x = 9$). Attractive women expressed preferences for all three clusters of men's mate characteristic (3Gs) (Buss and Shackelford, 2008), as the aforementioned conflict of good genes, good provider and good father, they have to trade off. Women placed more value on man's physical attractiveness, muscularity and immediate resource displays (Haselton and Gangestad, 2006) when pursuing short-term mating, in contrast, they placed greater importance on resource acquisition potential and good dad indicators when pursuing long-term mating (Buss and Shackelford, 2008; Lu et al., 2015). In the context of short-term sexual relationships, the perceived cues to high

parental investment in feminine men are of little value to women. Moreover, women of higher/lower physical condition can extract potential benefits from masculine men by copulating and conceiving with a short-term relationship. And therefore, both attractive and unattractive females trade on good genes (masculinity) in short-term context. On the contrary, in long-term relationships, better parenting and increased cooperation may outweigh the benefits of genetic fitness, thereby enhancing the attractiveness of more feminine males (Little and Mannion, 2006). Highly attractive women think that they are good enough to find another partner soon, and therefore, they do not need to be restricted by the real conditions or change their preferences according to the relationship context. However, women with poorer appearance have to trade off good genes and parental investment in the long-term context. As a result, the attractive-contingent masculine preference appeared more apparent in the long- than in the short-term context. A similar trend was observed in the voice domain (Feinberg et al., 2012).

On men's preference for women's femininity, the effect size for the short-term context was larger as compared to that for the long-term context (short-term = 0.202, $x = 2$; long-term = 0.081, $x = 3$), which indicated that attractive men were more strongly attracted by feminine female in short-term relationship than unattractive men, however, in long-term context, the men's self-attractiveness wasn't closely associated to their preference for femininity. Firstly, on the perspective of LH tradeoff strategies, preferences represent various LH strategies: fertility-related attributes (good genes) represent a fast LH strategy, whereas attributes of good parenting serve as slow LH function (Lu et al., 2017). Attractive men tended to adopt fast LH strategies with feminine women in short-term relationship, reported more short-term partners than less attractive men

TABLE 2 | Summary of the results of the sub-group meta-analysis.

Outcome	Moderator	Subgroup	X	Effect size (95% confidence interval)			Z	Q	I ² (%)	p
				Effect size(r)	lower	upper				
Masculinity	MOOA									0.000
		Objective	13	0.133	0.055	0.210	3.325***	28.868	58.431	
		Subjective	26	0.069	0.044	0.092	5.578***	41.651	39.977	
	Task									0.000
		Attractive rating	9	0.048	−0.061	0.157	0.864	19.591	59.165	
		FCT and AR	7	0.145	0.036	0.250	2.610**	7.438	19.332	
		FCT	12	0.081	0.033	0.129	3.281**	21.468	48.761	
		ST	11	0.183	0.065	0.297	3.018**	16.686	40.070	
	Stimulus									0.000
		Body	2	0.185	0.044	0.319	2.568**	0.126	0	
		Face	29	0.078	0.032	0.123	3.321***	54.769	48.877	
		Voice	8	0.184	0.079	0.285	3.407***	9.499	26.309	
	Same/opposite gender	Male stimulus								0.000 (same/opposite sex)
		Same	3	0.066	0.008	0.122	2.246	1.827	0.00	
		Opposite	35	0.107	0.059	0.155	4.302	71.000	52.112	
										0.028(RC)
Femininity	Task									0.000
		FCT and AR	4	−0.026	−0.238	0.188	−0.235	16.60	81.932	
		FCT	6	0.076	−0.039	0.188	1.300	42.069	88.115	
		ST	6	0.125	0.063	0.187	3.91***	5.835	14.315	
	Same/opposite gender	Female stimulus								0.005 (same/opposite sex)
		same	3	0.219	0.022	0.400	2.171*	17.6674**	88.676	
		opposite	7	0.083	0.008	0.156	2.179*	15.006***	60.016	
										0.000 (RC)

x, number of effect sizes; MOOA, Measures of Observers' Own Attractiveness; RC, Relationship Context; FCT and AR, Forced-choice Test and Attractiveness Rating; FCT, Forced-choice Test; ST, Sequence Test; ***p < 0.001, **p < 0.01, *p < 0.05.

(Rhodes et al., 2005). As without regards to raising the young, their attractiveness primed them with chance to produce greater number of offspring. Secondly, unlike women trading off between man's good gens and his resource, willing to invest, a man risked raising a child which is not his own, they mainly traded on partner's physical attractiveness and personality traits in long-term relationship (Little et al., 2014). Because feminine women are perceived as unfaithful and are considered more likely to have an affair or a brief sex trade. Both attractive and unattractive men considered women's risk of cuckoldry carefully in long-term contexts, such that self-attractiveness seemed less related with femininity preferences (Little et al., 2014).

The analysis of the moderating effect showed that the correlation coefficient for the relationship of the self-reported attractiveness of non-users of contraceptives and their masculinity preference was significantly higher than that of users (0.183 vs. 0.093). A cross-sectional longitudinal study confirmed that hormonal contraceptive users had a weaker preference for masculinity than non-users did (Little et al., 2013), which was consistent with the findings of other previous studies (Roberts et al., 2014). In addition, hormonal contraceptives tended to decrease women's overall physical attractiveness (Puts and Pope, 2013; Welling, 2013; Roberts et al., 2014), and as discussed above, women's attractiveness was their "market

TABLE 3 | The results of the meta-regression analysis.

Outcome variable	Moderator	<i>x</i>	<i>B</i>	<i>P</i>
Masculinity preference	Observer's age	30	−0.00467	0.016
	Published year	39	−0.01185	0.000
	Sample size of observers	39	−0.00004	0.009
Femininity preference	Observer's age	16	0.00222	0.163
	Published year	16	0.02641	0.004
	Sample size of observers	16	0.0004	0.402

value” and an important factor in “intra-sexual competition.” Therefore, by diminishing women’s attractiveness, hormonal contraceptives might make it more difficult for women to compete for romantic partners (Smith et al., 2009a; Puts and Pope, 2013). Similarly, effect size was more negatively correlated with observers’ age in masculine preference (see **Table 3**). The observers in the present analysis pool for masculine preference were females. In other words, younger attractive women preferred masculine faces. As people age, especially women, they recalibrate subjective impressions of their own attractiveness (i.e., impressions of their own “market value”), which, in turn, leads to a recalibration of their mate preferences (Little et al., 2011a). With a decrease in their market value with age, most women pay attention to parental characteristics (Jones et al., 2011).

Type of Stimuli

With reference to feminine preferences, because the present meta-analysis included only one study on vocal and none on body stimuli, we examined the moderating role of stimuli type in masculine preference. This analysis revealed significantly different effect sizes for body, voice, and face stimuli. Specifically, the effect sizes for voice and body were larger than that for face stimuli (body = 0.185, $x = 2$; voice = 0.184, $x = 8$; face = 0.078, $x = 29$), the effect sizes of body and voice are close to 0.2, a moderate amount. However, the coefficient of face and masculine preference was 0.078, a weak correlation, which was out of our expectation. Previous studies have interpreted covariation as evidence that different domains of masculinity all advertised a common underlying factor. Additionally, we suggested that men’s and women’s masculinity, as signaled by multiple traits, were related to some common information about the underlying quality of the observed individual. Of course, this did not mean that the signals overlap perfectly (Little et al., 2011a), and indeed, our data suggests that masculine preference in all the three traits was strongly related to observers’ self-attractiveness. This indicates a distinct characteristic. Masculine facial preferences are less relevant to self-attractiveness, potentially because a large number of studies have focused on facial masculine preference, while little attention has been paid to the preference for masculinity in the voice and body of the observed. Furthermore, most of these studies on facial preferences regarded self-attractiveness as a covariant, and they explored its interaction with other main variables (e. g., relationship context or menstrual cycle). Evidently, the correlation varied in different conditions, and therefore, more

effect sizes of facial masculine preference ($x = 29$) averaged into a smaller effect size.

The Task of Preference and Measures of Self-Attractiveness

Stronger effects were found when the sequence test (ST) was used instead of using the forced-choice test (FCT) with an attractiveness rating (Masculinity: ST = 0.183 $x = 11$; FCT = 0.081, $x = 12$; Femininity: ST = 0.125 $x = 6$; FCT = 0.076, $x = 6$). The sequence test provides observers a program in which they can regulate the sexual dimorphism independently until they reached the level that they considered most attractive (Carrito et al., 2016). The forced-choice test with an attractiveness rating merely provides observer two dimensions of a face (masculine and feminine) to choose from. They must then choose one that they consider more attractive and rate its attractiveness (Little and Mannion, 2006). Thus, the sequence test is more ecological and it reflects the observer’s preference for dimorphism more clearly.

Most studies included in the present meta-analysis focused on masculine preferences and none used an objective index to measure feminine preferences. Thus, we conducted this subgroup analysis only on masculine outcomes. A moderator analysis revealed a significant difference in the effect sizes of masculine preferences according to the measures of self-attractiveness, which is consistent with Penton-Voak et al.’s findings (Penton-Voak et al., 2003). They thought objective measures could be independent from menstrual cycle, and would be more stable than self-assessments are (Penton-Voak et al., 2003). Some researchers have emphasized that BMI and WHR were better indicators of female attractiveness (Swami et al., 2007; Smith et al., 2009b). Similarly, in the current meta-analysis, objective attractiveness was highly related to masculine preferences (effect size = 0.133, $x = 13$), but self-rated attractiveness showed a weak relationship (effect size = 0.069, $x = 26$). Most of the participants were female in the studies included in this meta-analysis. Subjective measures may be influenced by their individual physiology, for example, women perceived themselves more attractive when ovulating (Singh et al., 2001), and self-rated attractiveness may potentially fluctuate in a short space of time. Self-report measures are subjective and may not reflect how an individual is perceived by others. Therefore, objective measurements such as WHR, BMI, and other-rated attractiveness may be more sensitive to the relationship between self-attractiveness and sexual dimorphic preferences.

LIMITATIONS

We would like to mention some limitations of the present meta-analysis. The first concerns the number of studies analyzed. We included 25 studies, which were divided into sets based on feminine and masculine preferences before submitting them to separate meta-analyses. This meant that we had a limited number of studies for each moderator level (see **Table 2**). Therefore, some levels of the moderator were under-represented. The second

limitation of the current meta-analysis is that the age range of observers was relatively concentrated, in that most of them were teenagers or young individuals in their 20s, and only two studies selected older subjects with an average age of 33 years (Zietsch et al., 2015) and 48 years (Jones et al., 2011). Therefore, although the results of the present study revealed a significant regulatory effect of age, the accuracy of this correlation is difficult to prove.

CONCLUSIONS

This meta-analysis suggests that there is a real relationship between sexual dimorphic preferences and observers' self-attractiveness. Therefore, in future studies, researchers should control self-attractiveness to ensure the reliability of experimental results. According to the results of the present subgroup analyses, this relationship depends on relationship context, same/opposite-sex, and contraceptive using. These three moderating effects represented the observer's trade-off on 3Gs, were consistent with the life history strategies. Besides, measurement of observers' attractiveness, type of preference task and stimuli may also may involve the relationship. Therefore, in future studies, researchers should consider these factors and their interactions.

REFERENCES

- Bakker, T. C., Künzler, R., and Mazzi, D. (1999). Sexual selection: condition-related mate choice in sticklebacks. *Nature* 401, 234–234. doi: 10.1038/45727
- Boothroyd, L. G., Jones, B. C., Burt, D. M., DeBruine, L. M., and Perrett, D. I. (2008). Facial correlates of sociosexuality. *Evol. Hum. Behav.* 29, 211–218. doi: 10.1016/j.evolhumbehav.2007.12.009
- Boothroyd, L. G., Jones, B. C., Burt, D. M., and Perrett, D. I. (2007). Partner characteristics associated with masculinity, health and maturity in male faces. *Pers. Individ. Dif.* 43, 1161–1173. doi: 10.1016/j.paid.2007.03.008
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester: Wiley. doi: 10.1002/9780470743386
- Burke, D., and Sulikowski, D. (2010). A new viewpoint on the evolution of sexually dimorphic human faces. *Evol. Psychol.* 8, 573–585. doi: 10.1177/147470491000800404
- Burriss, R. P., Welling, L. L. M., and Puts, D. A. (2011). Men's attractiveness predicts their preference for female facial femininity when judging for short-term, but not long-term, partners. *Pers. Individ. Dif.* 50, 542–546. doi: 10.1016/j.paid.2010.11.022
- Buss, D. M., and Shackelford, T. K. (2008). Attractive women want it all: good genes, economic investment, parenting proclivities, and emotional commitment. *Evol. Psychol.* 6, 134–146. doi: 10.1177/147470490800600116
- Card, N. A. (2012). *Applied Meta-Analysis for Social Science Research*. New York, NY: Guilford Press.
- Carrito, M. L., dos Santos, I. M. B., Lefevre, C. E., Whitehead, R. D., da Silva, C. F., et al. (2016). The role of sexually dimorphic skin colour and shape in attractiveness of male faces. *Evol. Hum. Behav.* 37, 125–133. doi: 10.1016/j.evolhumbehav.2015.09.006
- Chang, L., Lu, H. J., and Zhu, X. Q. (2017). Good genes, good providers, and good fathers and mothers: the withholding of parental investment by married couples. *Evol. Behav. Sci.* 11, 199–211. doi: 10.1037/ebs0000086
- Chen, L., Jiang, J., Ren, Z., and Yuan, H. (2017). A meta-analysis of preference for sexual dimorphism of male's face. *Adv. Psychol. Sci.* 25, 553–569. doi: 10.3724/SP.J.1042.2017.00553
- Cornwell, R. E., Law Smith, M. J., Boothroyd, L. G., Moore, F. R., Davis, H. P., Stirrat, M., et al. (2006). Reproductive strategy, sexual development and

AUTHOR CONTRIBUTIONS

LC and ZR co-designed the study and wrote the manuscript. LC and YY conducted the literature searches, the studies selection. LC and XJ took charged. LC, XJ and HF analysed the data. LC and ZR wrote the manuscript. XJ and LC revised according the review.

FUNDING

This work was supported by the National Social Science Foundation of China (Grant No. CEA150173).

ACKNOWLEDGMENTS

We would like to acknowledge Dr. Robert P. Burriss (Department of psychology, Basel University, Basel, Switzerland) for his comments for the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02431/full#supplementary-material>

- attraction to facial characteristics. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 361, 2143–2154. doi: 10.1098/rstb.2006.1936
- Del Giudice, M., and Belsky, J. (2011). "The development of life history strategies: toward a multi-stage theory," in *The Evolution of Personality and Individual Differences*, eds D. M. Buss and P. H. Hawley (New York, NY: Oxford University Press), 154–176.
- Egger, M., Davey, S. G., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* 316, 469–471. doi: 10.1136/bmj.315.7109.629
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., and Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. *Evol. Hum. Behav.* 29, 233–241. doi: 10.1016/j.evolhumbehav.2007.12.008
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., Little, A. C., O'Connor, J. J. M., and Tigue, C. C. (2012). Women's self-perceived health and attractiveness predict their male vocal masculinity preferences in different directions across short- and long-term relationship contexts. *Behav. Ecol. Sociobiol.* 66, 413–418. doi: 10.1007/s00265-011-1287-y
- Fisher, M. L. (2004). Female intrasexual competition decreases female facial attractiveness. *Proc. Biol. Sci.* 271, S283–S285. doi: 10.1098/rsbl.2004.0160
- Fraccaro, P. J., Feinberg, D. R., DeBruine, L. M., Little, A. C., Watkins, C. D., and Jones, B. C. (2010). Correlated male preferences for femininity in female faces and voices. *Evol. Psychol.* 8, 447–461. doi: 10.1177/147470491000800311
- Gangestad, S. W., and Buss, D. M. (1993). Pathogen prevalence and human mate preferences. *Evol. Hum. Behav.* 14, 89–96.
- Gangestad, S. W., and Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *Behav. Brain Sci.* 23, 573–587. doi: 10.1017/S0140525X0000337X
- Haselton, M. (2005). The evolution of jealousy. *Trends Cognit. Sci.* 9, 506–507. doi: 10.1016/j.tics.2005.09.006
- Haselton, M. G., and Gangestad, S. W. (2006). Conditional expression of women's desires and men's mate guarding across the ovulatory cycle. *Horm. Behav.* 49, 509–518. doi: 10.1016/j.yhbeh.2005.10.006
- Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ* 327, 557–560. doi: 10.1136/bmj.327.7414.557

- Hittner, J. B., and Swickert, R. (2006). Sensation seeking and alcohol use: a meta-analytic review. *Addict. Behav.* 31, 1383–1401. doi: 10.1016/j.addbeh.2005.11.004
- Holzleitner, I. J., and Perrett, D. I. (2017). Women's preferences for men's facial masculinity: trade-off accounts revisited. *Adapt. Hum. Behav. Physiol.* 3, 304–320. doi: 10.1007/s40750-017-0070-3
- Jones, B. C., Boothroyd, L., Feinberg, D. R., and DeBruine, L. M. (2010). Age at menarche predicts individual differences in women's preferences for masculinized male voices in adulthood. *Pers. Individ. Diff.* 48, 860–863. doi: 10.1016/j.paid.2010.02.007
- Jones, B. C., DeBruine, L. M., Little, A. C., Conway, C. A., Welling, L. L. M., and Smith, F. (2007). Sensation seeking and men's face preferences. *Evol. Hum. Behav.* 28, 439–446. doi: 10.1016/j.evolhumbehav.2007.07.006
- Jones, B. C., Vukovic, J., Little, A. C., Roberts, S. C., and DeBruine, L. M. (2011). Circum-menopausal changes in women's preferences for sexually dimorphic shape cues in peer-aged faces. *Biol. Psychol.* 87, 453–455. doi: 10.1016/j.biopsycho.2011.04.004
- Kandrik, M. (2017). *Variation in Men's Mate Preferences and Mating Strategies*. Dissertation, University of Glasgow.
- Kandrik, M., and DeBruine, L. M. (2012). Self-rated attractiveness predicts preferences for opposite-sex faces, while self-rated sex-typicality predicts preferences for same-sex faces. *J. Evol. Psychol.* 10, 177–186. doi: 10.1556/JEP.10.2012.4.2
- Lefevre, C. E., and Saxton, T. K. (2017). Parental preferences for the facial traits of their offspring's partners can enhance parental inclusive fitness. *Evol. Hum. Behav.* 38, 546–551. doi: 10.1016/j.evolhumbehav.2017.01.006
- Lipsey, M. W., and Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage Publications.
- Little, A. C., Burriss, R. P., Petrie, M., Jones, B. C., and Roberts, S. C. (2013). Oral contraceptive use in women changes preferences for male facial masculinity and is associated with partner facial masculinity. *Psychoneuroendocrinology* 38, 1777–1785. doi: 10.1016/j.psyneuen.2013.02.014
- Little, A. C., Burt, D. M., Penton-Voak, I. S., and Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proc. Biol. Sci.* 268, 39–44. doi: 10.1098/rspb.2000.1327
- Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C., and Roberts, S. C. (2011a). Human preference for masculinity differs according to context in faces, bodies, voices, and smell. *Behav. Ecol.* 22, 862–868. doi: 10.1093/beheco/arr061
- Little, A. C., Jones, B. C., and Burriss, R. P. (2007). Preferences for masculinity in male bodies change across the menstrual cycle. *Horm. Behav.* 51, 633–639. doi: 10.1016/j.yhbeh.2007.03.006
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011b). Facial attractiveness: evolutionary based research. *Philos. Trans. Biol. Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Little, A. C., Jones, B. C., Feinberg, D. R., and Perrett, D. I. (2014). Men's strategic preferences for femininity in female faces. *Br. J. Psychol.* 105, 364–381. doi: 10.1111/bjop.12043
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. Biol. Sci.* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Little, A. C., and Mannion, H. (2006). Viewing attractive or unattractive same-sex individuals changes self-rated attractiveness and face preferences in women. *Anim. Behav.* 72, 981–987. doi: 10.1016/j.anbehav.2006.01.026
- Little, A. C., Saxton, T. K., Roberts, S. C., Jones, B. C., DeBruine, L. M., Vukovic, J., et al. (2010). Women's preferences for masculinity in male faces are highest during reproductive age range and lower around puberty and post-menopause. *Psychoneuroendocrinology* 35, 912–920. doi: 10.1016/j.psyneuen.2009.12.006
- Lu, H. J., Wong, K. C., and Chang, L. (2017). The association between life history strategy and mate preference in men. *Pers. Individ. Diff.* 116, 157–163. doi: 10.1016/j.paid.2017.04.047
- Lu, H. J., Zhu, X. Q., and Chang, L. (2015). Good genes, good providers, and good fathers: economic development involved in how women select a mate. *Evol. Behav. Sci.* 9, 215–288. doi: 10.1037/ebs0000048
- Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int. J. Surg.* 8, 336–341. doi: 10.1016/j.ijsu.2010.02.007
- Moore, F. R., Law Smith, M. J., Taylor, V., and Perrett, D. I. (2011). Sexual dimorphism in the female face is a cue to health and social status but not age. *Pers. Individ. Diff.* 50, 1068–1073. doi: 10.1016/j.paid.2011.01.026
- O'Connor, J. J. M., Feinberg, D. R., Fraccaro, P. J., Borak, D. J., Tigue, C. C., Jones, B. C., et al. (2012). Female preferences for male vocal and facial masculinity in videos. *Ethology* 118, 321–330. doi: 10.1111/j.1439-0310.2011.02013.x
- Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P., and Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*). *J. Comp. Psychol.* 117, 264–271. doi: 10.1037/0735-7036.117.3.264
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* 26, 388–397. doi: 10.1016/j.evolhumbehav.2005.03.001
- Puts, D. A., and Pope, L. E. (2013). Moderators, mates, and matchmakers: effects of oral contraceptives on sexual desire may also depend on partners' behavior and the role of female choice. *Arch. Sex. Behav.* 42, 1379–1380. doi: 10.1007/s10508-013-0165-2
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Rhodes, G., Chan, J., Zebrowitz, L. A., and Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proc. Biol. Sci.* 270, S93–S95. doi: 10.1098/rsbl.2003.0023
- Rhodes, G., Simmons, L. W., and Peters, M. (2005). Attractiveness and sexual behavior: does attractiveness enhance mating success? *Evol. Hum. Behav.* 26, 186–201. doi: 10.1016/j.evolhumbehav.2004.08.014
- Roberts, S. C., Little, A. C., Burriss, R. P., Cobey, K. D., Klapilová, K., Havlíček, J., et al. (2014). Partner choice, relationship satisfaction, and oral contraception: the congruency hypothesis. *Psychol. Sci.* 25, 1497–1503. doi: 10.1177/0956797614532295
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychol. Bull.* 86, 638–641. doi: 10.1037/0033-2909.86.3.638
- Singh, D., Davis, M., and Randall, P. (2001). "Flaunting ovulation: lower WHR, enhanced self-perceived attractiveness, and increased sexual desire," *Paper presented at the 13th Human Behavior and Evolution Society Meeting* (London: University College London).
- Smith, F. G., Jones, B. C., Little, A. C., DeBruine, L. M., Welling, L. L. M., Vukovic, J., et al. (2009a). Hormonal contraceptive use and perceptions of trust modulate the effect of relationship context on women's preferences for sexual dimorphism in male face shape. *J. Evol. Psychol.* 7, 195–210. doi: 10.1556/JEP.7.2009.3.1
- Smith, F. G., Jones, B. C., Welling, L. L. W., Little, A. C., Vukovic, J., Main, J. C., et al. (2009b). Waist-hip ratio predicts women's preferences for masculine male faces, but not perceptions of men's trustworthiness. *Pers. Individ. Diff.* 47, 476–480. doi: 10.1016/j.paid.2009.04.022
- Swami, V., Smith, J., Tsiokris, A., Georgiades, C., Sangareau, Y., Tovee, M. J., et al. (2007). Male physical attractiveness in Britain and Greece: a cross-cultural study. *J. Soc. Psychol.* 147, 15–26. doi: 10.3200/SOCP.147.1.15-26
- Thornhill, R., and Gangestad, S. W. (1993). Human facial beauty: averageness, symmetry, and parasite resistance. *Hum. Nat.* 4, 237–269. doi: 10.1007/BF02692201
- Thornhill, R., and Gangestad, S. W. (1996). The evolution of human sexuality. *Trends Ecol. Evol.* 11, 98–102. doi: 10.1016/0169-5347(96)81051-2
- Thornhill, R., and Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evol. Hum. Behav.* 27, 131–144. doi: 10.1016/j.evolhumbehav.2005.06.001
- Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L., Little, A. C., et al. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Pers. Individ. Diff.* 45, 451–456. doi: 10.1016/j.paid.2008.05.013
- Vukovic, J., Jones, B. C., DeBruine, L., Feinberg, D. R., Smith, F. G., Little, A. C., et al. (2010). Women's own voice pitch predicts their preferences for masculinity in men's voices. *Behav. Ecol.* 21, 767–772. doi: 10.1093/beheco/arq051

- Welling, L. L. (2013). Psychobehavioral effects of hormonal contraceptive use. *Evol. Psychol.* 11, 718–742. doi: 10.1177/147470491301100315
- Welling, L. L. M., DeBruine, L. M., Little, A. C., and Jones, B. C. (2009). Extraversion predicts individual differences in women's face preferences. *Pers. Individ. Dif.* 47, 996–998. doi: 10.1016/j.paid.2009.06.030
- Welling, L. L. M., Jones, B. C., and DeBruine, L. M. (2008). Sex drive is positively associated with women's preferences for sexual dimorphism in men's and women's faces. *Pers. Individ. Dif.* 44, 161–170. doi: 10.1016/j.paid.2007.07.026
- Widemo, F., and Sæther, S. A. (1999). Beauty is in the eye of the beholder: causes and consequences of variation in mating preferences. *Trends Ecol. Evol.* 14, 26–31. doi: 10.1016/S0169-5347(98)01531-6
- Wood, W., Kressel, L., Joshi, P. D., and Louie, B. (2014). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Rev.* 6, 229–249. doi: 10.1177/1754073914523073
- Zheng, L. J., and Zheng, Y. (2016). Preferences for masculinity across faces, bodies, and personality traits in homosexual and bisexual chinese men: relationship to sexual self-labels and attitudes toward masculinity. *Arch. Sex. Behav.* 45, 725–733. doi: 10.1007/s10508-015-0543-z
- Zietsch, B. P., Lee, A. J., Sherlock, J. M., and Jern, P. (2015). Variation in women's preferences regarding male facial masculinity is better explained by genetic differences than by previously identified context-dependent effects. *Psychol. Sci.* 26, 1440–1448. doi: 10.1177/0956797615591770

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Chen, Jiang, Fan, Yang and Ren. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Relative Contribution of Jawbone and Cheekbone Prominence, Eyebrow Thickness, Eye Size, and Face Length to Evaluations of Facial Masculinity and Attractiveness: A Conjoint Data-Driven Approach

Justin K. Mogilski^{1*} and Lisa L. M. Welling²

OPEN ACCESS

Edited by:

Achim Schuetzwohl,
Brunel University London,
United Kingdom

Reviewed by:

R. Nathan Pipitone,
Florida Gulf Coast University,
United States
Darren Burke,
University of Newcastle, Australia

*Correspondence:

Justin K. Mogilski
jmogilsk@mailbox.sc.edu;
justin.mogilski@gmail.com

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 10 September 2018

Accepted: 19 November 2018

Published: 05 December 2018

Citation:

Mogilski JK and Welling LLM
(2018) The Relative Contribution
of Jawbone and Cheekbone
Prominence, Eyebrow Thickness, Eye
Size, and Face Length to Evaluations
of Facial Masculinity
and Attractiveness: A Conjoint
Data-Driven Approach.
Front. Psychol. 9:2428.
doi: 10.3389/fpsyg.2018.02428

¹ Department of Psychology, University of South Carolina Salkehatchie, Walterboro, SC, United States, ² Department of Psychology, Oakland University, Rochester, MI, United States

Recent work demonstrates the methodological rigor of a type of data-driven analysis (i.e., conjoint analysis; CA), which accounts for the relative contribution of different facial morphological cues to interpersonal perceptions of romantic partner quality. This study extends this literature by using a conjoint face ranking task to predict the relative contribution of five sexually dimorphic facial shape features (jawbone and cheekbone prominence, eyebrow thickness, eye size, face length) to participants' ($N = 922$) perceptions of facial attractiveness and sex-typicality (i.e., masculinity/femininity). For overall partner attractiveness, eyebrow thickness and jawbone prominence were relatively more salient than cheekbone prominence and eye size. Interestingly, masculinized (i.e., thicker) eyebrows were marginally more attractive for female than male faces, particularly within a long-term mating context. Masculinized jawbone prominence was more attractive for male than female faces, and feminized jawbone prominence was more attractive for female than male faces. For perceptions of masculinity, eyebrow thickness, jawbone prominence, and facial height were relatively more salient than cheekbone prominence and eye size, although facial height was more important for female than male faces, and jawbone prominence was marginally more important for male than female faces. These findings highlight the prominence of eyebrows, the jawline, and facial height during perception of facial attractiveness and masculinity – though it should be noted that many of these differences were small to moderate in effect size. Findings are interpreted in the context of prior research, and future directions for studying why these facial traits exhibit superior signaling capacity are discussed.

Keywords: face preference, face shape, masculinity, attractiveness, conjoint analysis, data-driven, digital manipulation

INTRODUCTION

Facial morphological cues (e.g., shape, color, and texture) are indicators of underlying physiology (Stephen et al., 2009; Little et al., 2011b; Jones et al., 2012). From these cues, humans can accurately predict certain physical and psychological qualities (e.g., an individual's health, physical attractiveness, trustworthiness) that are significant to partner selection and social judgment (e.g., Zebrowitz, 2011; Todorov et al., 2015). These qualities can be assessed by observers from facial photographs at first acquaintance, and significantly impact employment decisions, mate selection, friendship, and other key aspects of social interaction (e.g., Petrican et al., 2014; Walker and Vetter, 2016; Funk et al., 2017). Face perception researchers have studied how these cues are processed during interpersonal evaluation by digitally manipulating photographic facial cues and presenting these images to third-party raters. These manipulations predictably alter perceptions of attractiveness, dominance, sex-typicality (i.e., masculinity/femininity), health, trustworthiness, and other social attributes (for a review, see Todorov et al., 2008; Little et al., 2011a). However, this research has tended to focus on individual facial cues in isolation, and debate is now turning to the relative contributions of these cues to social perception (Scott et al., 2010; Stephen et al., 2012; Mogilski and Welling, 2017).

Data-driven models are becoming more valued within scientific face perception research for their capacity to account for a broader array of structural and configural facial features, and the contributions of those specific features to social perception, compared to traditional methods alone (Todorov et al., 2011; Zhang et al., 2018). Early studies that used digital face stimuli to alter and assess person perception (e.g., Perrett et al., 1999; Penton-Voak et al., 2001) measure the influence of perceptually distinct facial cues (e.g., symmetry and sexual dimorphism) by digitally altering one feature while experimentally or statistically controlling for variation in other features. Although this work makes compelling contributions to the literature by reducing confounds, it tells us little about how collections of traits are evaluated in combination. Recent data-driven techniques have overcome this limitation by permitting distinct clusters of features to be altered simultaneously (e.g., Mogilski and Welling, 2017; Stephen et al., 2017; Jones, 2018). For example, Stephen et al. (2017) recorded participants' physiological health (i.e., blood pressure, BMI, percent body fat) and regressed these measures onto variation in facial morphology. These measures were then subjected to factor analysis to identify which clusters of facial features predict variation in health. Participants were asked to change the appearance of potential romantic partners' facial photographs to fit their preference using sliders that incrementally altered faces along each health dimension, thereby digitally manipulating the constellation of facial features naturally associated with variation in these health indices. Similarly, Jones (2018) recorded participants' self-reported health and implemented a Brunswick lens model to assess which facial cues are utilized to assess health, and which cues are valid indicators of health. Photographs were subsequently manipulated according to whichever cues were most utilized and valid.

Compared to prior techniques, these methods identify facial cues and manipulate them via digital transformations that are based on a broad collection of naturalistic features, rather than from artificially restricted parameters based on theory alone.

Although these models have been particularly useful for initially exploring, identifying, and simulating the facial cues that contribute to person perception and social decision-making, they are limited in their capacity to assess the relative contribution of several concurrently altered facial cues to holistic perception of those faces. In the studies noted above (Stephen et al., 2017; Jones, 2018), digital transformations were applied to facial images and rated sequentially rather than concurrently. For example, Stephen et al. (2017) asked participants to manipulate a series of faces to appear as healthy as possible by manipulating apparent BMI, blood pressure, and body fat, but each of these dimensions was manipulated independently and rated across separate trials. Jones (2018) manipulated two features concurrently (i.e., averageness and color), but asked participants to assess stimuli of different combinations (e.g., high averageness, low color) across separate line-ups of faces. These methods allow researchers to examine preference for feature combinations, but they are limited by how many feature combinations may be examined at the same time without separating them into separate trials or experimental conditions.

Conjoint Analysis

Conjoint analysis (CA) provides a convenient way to overcome this design challenge. CA is a multivariate, data-driven analysis used in marketing research (e.g., Gustafsson et al., 2007) that has recently been adapted to study human mate preferences (Mogilski et al., 2014; Mogilski and Welling, 2017). Generally, CA is used to assess how individuals make trade-offs among multiple attributes when evaluating "whole" units that comprise those attributes. For example, CA is often used to evaluate which attributes of a product are most important during consumer purchasing decisions by having consumers rank several versions of the product, where each version is composed of a unique combination of product attributes. Mogilski and Welling (2017) first used this technique to examine the relative salience of three facial cues (i.e., sexual dimorphism, color, and symmetry) during romantic partner perception. Compared to other methods, this technique allows researchers to present sets of faces wherein each face is altered by several different features at once. Participants rank these sets on some metric (e.g., their attractiveness as a romantic partner) and CA provides measures of the relative contribution of each feature to participants' overall ranking decisions. Using this technique, Mogilski and Welling (2017) found that facial shape masculinity/femininity was relatively more important than both symmetry and color cues to health during participants' evaluations of potential romantic partners' facial photographs. However, presenting individuals with multiple versions of mates who vary across several different traits is but one potential use of CA. This technique can also be used to explore preferences for feature variations that relate to a single construct.

Specifically, CA can investigate how several traits that contribute toward a single construct impact rater's perceptions of that construct.

Current Study

The present study contributes to current face perception literature by using CA to assess the relative contributions of several facial shape cues to perceptions of romantic partner attractiveness and masculinity. Previous research (e.g., Scheib et al., 1999; Penton-Voak et al., 2001; Koehler et al., 2004; Thornhill and Gangestad, 2006; Apicella et al., 2008) has identified several prominent shape cues that contribute to perceptions of sexual dimorphism (i.e., features that differ statistically between male and female faces): eyebrow prominence, cheekbone prominence, eye size, facial height, and jawbone prominence. Among these features, eyebrow prominence, facial height, and jawbone prominence are all reliably larger in men (i.e., are positively related to a masculine face shape and negatively related to a feminine face shape), whereas cheekbone prominence and eye size were reliably larger in women (i.e., are positively related to a feminine face shape and negatively related to a masculine face shape). Identification of the specific features that impact the overall measured sexual dimorphism of a particular individual is important for perceptual research, but no research to date has investigated how these individual features are prioritized relative to one another with respect to mate choice or with respect to the overall evaluation of a person's masculinity/femininity.

This study assessed whether individuals prioritize certain sexually dimorphic facial cues of partner quality (i.e., eyebrow prominence, cheekbone prominence, eye size, facial height, and jawbone prominence) when evaluating the attractiveness of same- and opposite-sex individuals' facial photographs as long- and short-term romantic partners. Furthermore, it investigated whether individuals prioritize any of these specific features when ranking different versions of the same individual by perceived masculinity. Given that face perception is, in part, due to part-based information processing (Schwaninger et al., 2004; McKone and Yovel, 2009), it is possible that the part-worth value of some facial features is weighted more heavily compared to others during partner perception. Similarly, some features may be more salient than others within different mating contexts (e.g., long- versus short-term; Buss and Schmitt, 1993; Gangestad and Simpson, 2000). Examining how these features are prioritized within long- (i.e., committed) versus short-term (i.e., purely sexual) mating contexts may explain the specific signal value of the distinct facial shape cues that contribute to person perception, and thereby reveal which features are most important to perceptions of attractive facial cues (e.g., Thornhill and Gangestad, 1999; Gangestad and Scheyd, 2005; Puts, 2010). Indeed, developing techniques that improve the accuracy of models that estimate psychological and physiological qualities from facial information (see Hu et al., 2017; Todorov, 2017) is a critical future direction in face perception research (see Jack and Schyns, 2017, for a review). Moreover, because this is the first study to manipulate individual features rather than whole faces by masculinity and femininity, the contribution of

each manipulation to perceptions of facial masculinity will reveal interesting information about how we process this trait, as well as serve to further verify the computer graphics manipulations. Finally, this study will expand prior findings that used CA to study face perception (Mogilski and Welling, 2017). Because this study found that sexually dimorphic shape cues were more important than color and symmetry cues (Cohen's $d \sim 0.60$), this study sought to examine which shape features were driving this effect. That is, this study examines whether some shape features signal relatively more information about partner quality (i.e., attractiveness; masculinity/femininity) than others.

MATERIALS AND METHODS

Participants

Participants ($N = 922$, 250 male; age: $M = 20.22$ years, $SD = 3.53$; range = 18–51) were recruited from a university in the mid-western United States and various social media outlets (e.g., Facebook, Reddit, Twitter). The majority of participants were White (78.4%; Black 9.5%, Asian 5.5%, Hispanic/Latino 2.3%, "Other" 4.3%), roughly half reported currently being single (49% versus 51% reported being in a romantic relationship), and the majority reported being exclusively heterosexual (91.3%; 6.5% bisexual, 2.2% exclusively homosexual).

Stimuli

Using well-established methods (e.g., Jones et al., 2005; Little et al., 2007; Welling et al., 2008), composite male and female faces were generated by averaging the shape, color, and texture of a group of 60 Caucasian adult male faces and a group of 60 Caucasian adult female faces. Each composite served as the base image for a set of 19 photographs that varied exclusively by a series of objective, composite-based image transformations (detailed below). Up to five distinct facial characteristics were transformed per photograph variation: eyebrow prominence, cheekbone prominence, eye size, facial height, and jawbone prominence. These features are sexually dimorphic and vary with perceptions of facial attractiveness (Keating, 1985; Scheib et al., 1999; Penton-Voak et al., 2001; Baudouin and Tiberghien, 2004). To permit CA of participants' photograph rankings, each of the 19 photograph variations were planned using an orthogonal array generated with IBM SPSS 21, which is constructed according to a standard formula drawn from statistical reference material. A fractional-factorial design was used to minimize the number of photograph variations that participants were required to rank (Hair et al., 1995). This design generates the fewest number of profiles needed to estimate the contribution of each of the five facial characteristics to overall face evaluation. Each of the five facial features were assigned three possible levels (i.e., feature masculinization, unaltered, or feature feminization), indicating which transformations would be applied to each photograph. This produced an orthogonal array of 16 photograph variations, whereby each variation possessed a unique combination of the five facial characteristics. For example, a face might have masculinized eyebrow prominence, feminized cheekbone prominence, unaltered eye size, feminized facial

height, and masculine jawbone prominence. Three additional holdout images (for a total of 19 photographs) were included to test the validity of the CA utility estimates (see Hair et al., 1995; utility estimates and holdout profiles are defined in more detail in the Results section below). All participants ranked the same 19 images constructed based on the orthogonal array.

To alter specific facial features, image transformation methods used in prior work (e.g., DeBruine et al., 2006; Welling et al., 2007, Welling et al., 2008) were adapted to target individual features rather than whole faces. Specifically, rather than apply composite-based transformations holistically to base images (i.e., to the whole face), 5 male and 5 female composite images were first created, whereby each composite image possessed one feature (i.e., eyebrow prominence, cheekbone prominence, eye size, facial height, or jawbone prominence) of the opposite-sex, but that was otherwise sex-typical. Thus, for each composite image, the points that correspond to individual features (e.g., eye size) were altered such that they matched the position of those same points on the opposite-sex composite (see **Figure 1** for an example). These composite images were then applied to same-sex base images by taking 50% of the linear differences in 2D shape between the applicable altered composites (e.g., **Figure 1**, image C) and the original same-sex composites (i.e., **Figure 1**, image A for women, image B for men) and adding to or subtracting from corresponding points on the base image. **Figures 2, 3** demonstrate the complete array of masculine and feminine manipulations individually applied to base female (**Figure 2**) and male (**Figure 3**) composite faces. These transformations were then applied to base images (i.e., the original, unaltered composite images) according to the orthogonal array (see **Table 1**). In other words, facial features were manipulated as per previous research (e.g., Welling et al., 2007, Welling et al., 2008) except that individual features were independently manipulated and then concurrently applied to the same face (see examples in **Figure 4**). Although no study has manipulated individual facial features in this way, holistic facial manipulations using these techniques have been shown to influence perceptions of masculinity and femininity in the predicted directions (DeBruine et al., 2006; Welling et al., 2007).

Procedure

All experimental materials were presented using Qualtrics. After indicating their consent, participants provided demographic information (i.e., age, sex, ethnicity, relationship status, and sexual orientation) and then completed a series of four face ranking tasks. For each task, participants were presented with each of two sets (one male, one female) of nineteen digital facial photographs. Participants were asked to rank the images within each set relative to one another twice: once according to their preference for a long-term relationship and once according to their preference for a short-term relationship. Long- and short-term relationships were defined for participants as follows:

Long-term relationship: You are looking for the type of person who would be attractive in a long-term relationship. Examples of this type of relationship would include someone you may want to move in with, someone you may consider leaving a current partner to be

with, and someone you may, at some point, wish to marry (or enter into a relationship on similar grounds as marriage).

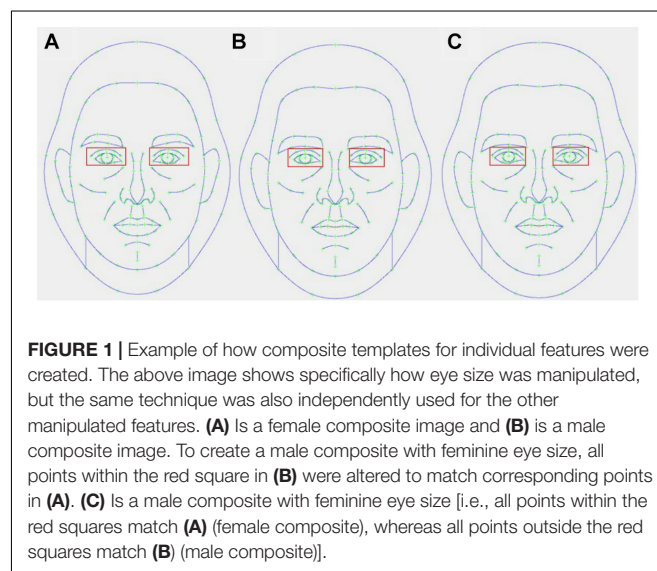
Short-term relationship: You are looking for the type of person who would be attractive in a short-term relationship. This implies that the relationship may not last a long time. Examples of this type of relationship would include a single date accepted on the spur of the moment, an affair within a long-term relationship, and the possibility of a one-night stand.

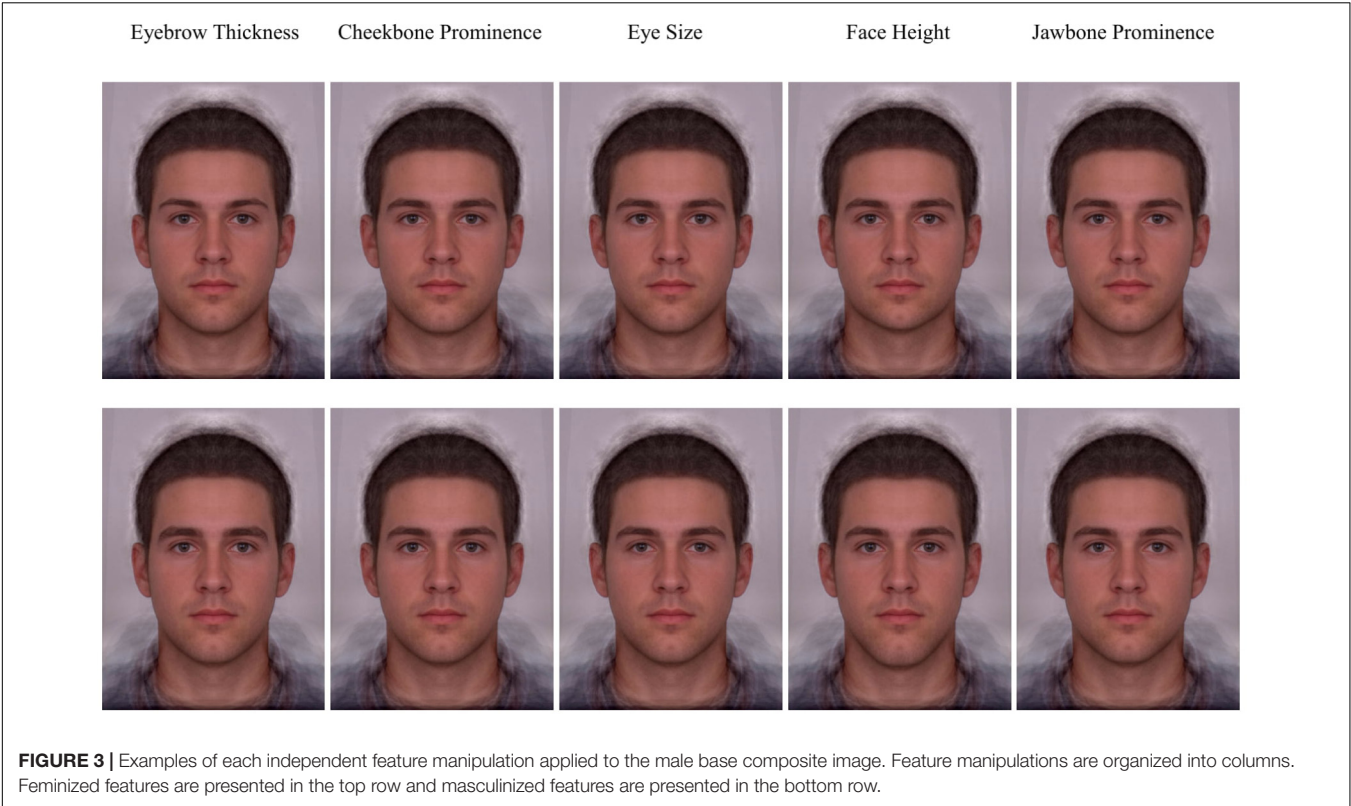
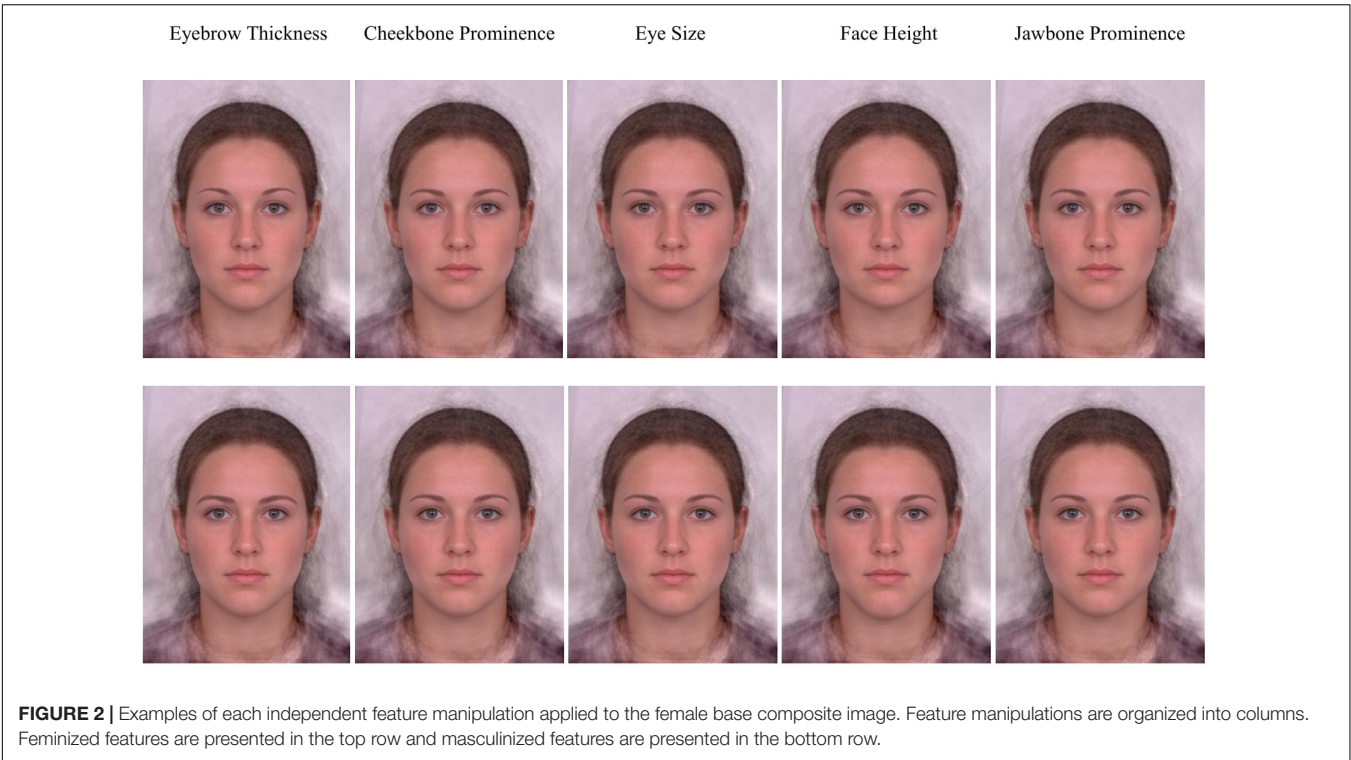
Participants were instructed to rank same-sex photographs according to how they believed a heterosexual person of the opposite-sex would rank them. The order in which the face ranking tasks and photographs within sets were presented was randomized.

RESULTS

CA was performed to assess the relative importance of each of the five facial features in participants' ranking decisions. CA produces importance values, which indicate a feature's overall contribution to how profiles are ranked (e.g., the overall importance of cheekbone prominence, eye size, etc.), and part-worth utility estimates, which indicate the relative importance of each level within each trait (i.e., masculinization, feminization, unaltered). In other words, importance values reveal which features are weighted most heavily relative to others during ranking decisions, but not the direction of preference within any given feature (e.g., whether eyebrow prominence is more important for ratings of attractiveness than cheekbone prominence, but not whether masculine, unaltered, or feminine eyebrow prominence is preferred). On the other hand, utility estimates reveal the importance of the manipulation within a trait (e.g., preference for masculine, unaltered, versus feminine eyebrow prominence). Importance values and part-worth utility estimates were calculated for each set of faces.

Participants' rankings of holdout profiles were accurately predicted by the utility estimates (all $\tau = 1.00$) for both





attractiveness and masculinity assessments. Holdout profiles are image variations with unique facial characteristic combinations that are ranked alongside the original 16 profile, but which

are not used to generate importance values or utility estimates. The attractiveness or masculinity of a holdout profile can be calculated using the model generated from

TABLE 1 | Orthogonal array of facial transformations and the respective image variations to which they were applied.

Image variation	Eyebrow	Cheekbone	Eye Size	Face Height	Jawbone
1	Unaltered	Masculine	Feminine	Masculine	Feminine
2	Unaltered	Feminine	Masculine	Masculine	Masculine
3	Masculine	Feminine	Unaltered	Masculine	Unaltered
4	Masculine	Unaltered	Masculine	Feminine	Feminine
5	Masculine	Masculine	Feminine	Unaltered	Masculine
6	Feminine	Masculine	Masculine	Unaltered	Unaltered
7	Feminine	Unaltered	Masculine	Masculine	Masculine
8	Masculine	Unaltered	Feminine	Masculine	Unaltered
9	Unaltered	Masculine	Masculine	Feminine	Unaltered
10	Feminine	Feminine	Feminine	Feminine	Masculine
11	Masculine	Masculine	Unaltered	Feminine	Masculine
12	Feminine	Masculine	Unaltered	Unaltered	Masculine
13	Unaltered	Unaltered	Unaltered	Unaltered	Masculine
14	Masculine	Feminine	Masculine	Unaltered	Feminine
15	Masculine	Masculine	Masculine	Masculine	Masculine
16	Masculine	Masculine	Masculine	Masculine	Masculine
17 (Holdout)	Masculine	Feminine	Masculine	Masculine	Unaltered
18 (Holdout)	Masculine	Unaltered	Unaltered	Masculine	Feminine
19 (Holdout)	Masculine	Masculine	Masculine	Masculine	Feminine

The number listed in the "Image variation" column corresponds to each image's numerical label in **Figure 4**.

participants' rankings of the other 16 images. That is, the utility estimates for each characteristic (e.g., masculinized face height, feminized eye size, etc.) can be summed within each image variation to give an overall estimated attractiveness or masculinity score for that image. How well the score for each holdout profile predicts participants' rankings of those profiles relative to each of the other 16 profiles is represented by the tau coefficient. In other words, tau represents of how well the utility estimates predict participants' rankings of the holdout profiles relative to the other 16 profiles.

For each of the following analyses, participant gender was included as an additional variable. All interactions with gender were nonsignificant even before Bonferroni correction (adjusted critical $p = 0.01$; all $ps > 0.08$). Therefore, participant gender was excluded from our report below.

Attractiveness Ratings Importance Values

looseness2 A $2(\text{sex of face [male, female]}) \times 2(\text{relationship context [long-term, short-term]}) \times 5(\text{facial attribute [eyebrow thickness, cheekbone prominence, eye size, face height, jawbone prominence]})$ repeated-measures ANOVA was used to examine differences in importance values for each facial attribute in male and female faces ranked for desirability as long- and short-term mates. All *post hoc* analyses and pairwise comparisons were adjusted using Bonferroni correction (critical $p = 0.01$). There was a main effect for facial attribute, $F(4, 3684) = 23.41$, $p < 0.001$, $\eta^2 = 0.03$. Importance values for eyebrow thickness ($M = 21.28$, $SD = 6.62$) were not significantly different from face height ($M = 20.31$, $SD = 7.61$, $p = 0.114$) or jawbone prominence

($M = 20.48$, $SD = 5.66$, $p = 0.155$), but both eyebrow thickness and jawbone prominence were greater than cheekbone prominence ($M = 18.43$, $SD = 4.42$, $p < 0.001$, $d = 0.31$, $d = 0.29$) and eye size ($M = 19.49$, $SD = 5.36$, $p < 0.001$, $d = 0.19$, $d = 0.12$) when ranking faces for attractiveness. Likewise, importance values were greater for face height than for cheekbone prominence ($p < 0.001$, $d = 0.19$).

There was also a significant interaction between sex of face and facial attribute, $F(4, 3684) = 6.97$, $p < 0.001$, $\eta^2 = 0.01$. Importance values for eye size were greater for male faces ($M = 20.09$, $SD = 7.28$) than for female faces ($M = 18.90$, $SD = 7.12$), $t(921) = 3.76$, $p < 0.001$, $d = 0.12$, whereas importance values for face height were greater for female faces ($M = 20.90$, $SD = 9.98$) than for male faces ($M = 19.73$, $SD = 8.36$), $t(921) = 3.45$, $p = 0.001$, $d = 0.12$. There was also a significant interaction between relationship context and facial attribute, $F(4, 3684) = 2.39$, $p = 0.049$, $\eta^2 = 0.003$. Importance values for eyebrow thickness were greater for a long-term ($M = 21.71$, $SD = 8.26$) compared to short-term ($M = 20.95$, $SD = 8.17$) relationship context, however, this was not significant after Bonferroni correction, $t(921) = 2.04$, $p = 0.042$, $d = 0.08$.

Utility Estimates

Five $2(\text{sex of face [male, female]}) \times 2(\text{relationship context [long-term, short-term]}) \times 3(\text{attribute level [masculinized, unaltered, feminized]})$ repeated-measures ANOVAs were used to examine differences in utility estimates for each level of each facial attribute in male and female faces ranked for desirability as long- and short-term mates. All *post hoc* analyses and pairwise comparisons were adjusted using Bonferroni correction (critical $p = 0.017$).

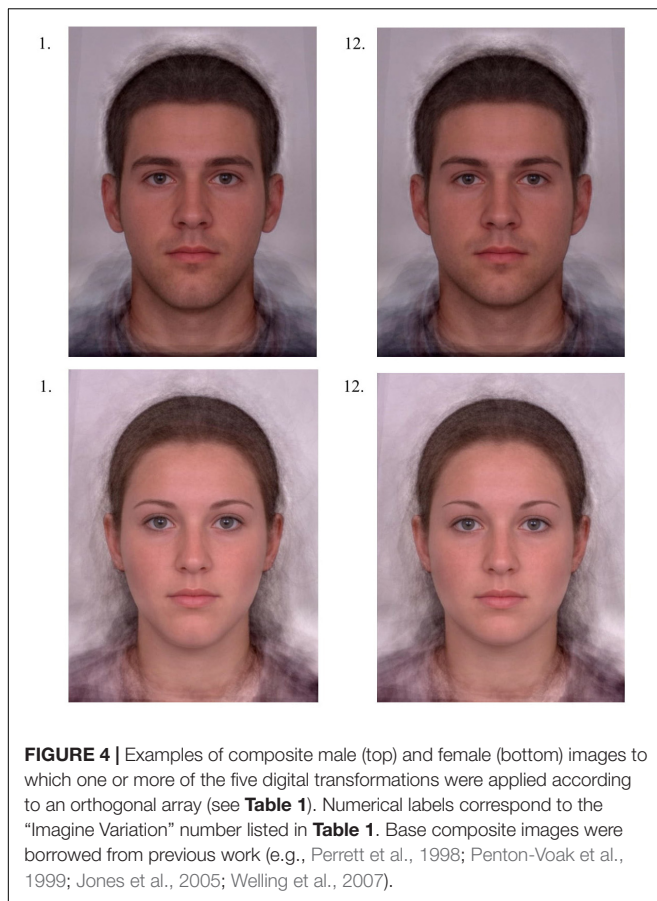


FIGURE 4 | Examples of composite male (top) and female (bottom) images to which one or more of the five digital transformations were applied according to an orthogonal array (see **Table 1**). Numerical labels correspond to the “Imagine Variation” number listed in **Table 1**. Base composite images were borrowed from previous work (e.g., Perrett et al., 1998; Penton-Voak et al., 1999; Jones et al., 2005; Welling et al., 2007).

Eyebrow thickness

There was a main effect for attribute level, $F(2, 1842) = 27.37$, $p < 0.001$, $\eta^2 = 0.03$. Utility estimates were greater for masculinized ($M = 0.25$, $SD = 0.95$) than for unaltered ($M = -0.12$, $SD = 1.03$, $p < 0.001$, $d = 0.23$) and feminized ($M = -0.13$, $SD = 1.13$, $p < 0.001$, $d = 0.21$) eyebrow thickness. This was moderated by a significant interaction with sex of face, $F(2, 1842) = 3.81$, $p = 0.022$, $\eta^2 = 0.004$. Utility estimates for masculinized eyebrow thickness were greater for women ($M = 0.33$, $SD = 1.35$) than for men ($M = 0.17$, $SD = 1.19$), $t(921) = 2.85$, $p = 0.004$, $d = 0.10$. This was further moderated by a three-way interaction, $F(2, 1842) = 3.14$, $p = 0.044$, $\eta^2 = 0.003$. For female faces, utility estimates for masculinized eyebrows were higher in a long- ($M = 0.42$, $SD = 1.75$) versus short-term ($M = 0.25$, $SD = 1.68$) context, $t(921) = 2.39$, $p = 0.017$, $d = 0.11$, though this was only marginally significant after Bonferroni correction. There were no other significant main effects or interactions (all $F < 1.96$, all $p > 0.141$).

Cheekbone prominence

There was a main effect for attribute level, $F(2, 1842) = 30.63$, $p < 0.001$, $\eta^2 = 0.03$. Utility estimates were greater for masculinized ($M = 0.21$, $SD = 0.79$) than unaltered ($M = -0.11$, $SD = 0.78$, $p < 0.001$, $d = 0.25$) and feminized ($M = -0.11$,

$SD = 0.90$, $p < 0.001$, $d = 0.21$) cheekbones. There were no other significant main effects or interactions (all $F < 2.09$, all $p > 0.124$).

Eye size

There was a main effect for attribute level, $F(2, 1842) = 58.21$, $p < 0.001$, $\eta^2 = 0.06$. Utility estimates were higher for unaltered ($M = 0.23$, $SD = 0.78$) than masculinized ($M = 0.09$, $SD = 0.91$, $p = 0.006$, $d = 0.11$) and feminized ($M = -0.32$, $SD = 1.07$, $p < 0.001$, $d = 0.30$) eye size. Estimates were also higher for masculinized compared to feminized eye size ($p < 0.001$, $d = 0.25$). There was also a significant interaction between relationship context and attribute level, $F(2, 1842) = 3.37$, $p = 0.035$, $\eta^2 = 0.004$. Utility estimates for masculinized eye size were greater in a short- ($M = 0.15$, $SD = 1.04$) versus long-term ($M = 0.04$, $SD = 1.05$) context, $t(921) = -2.37$, $p = 0.018$, $d = 0.08$, whereas unaltered eye size was preferred in a long- ($M = 0.29$, $SD = 1.21$) compared to short-term ($M = 0.16$, $SD = 1.24$) context, $t(921) = 2.33$, $p = 0.020$, $d = 0.01$, however, both were nonsignificant after Bonferroni correction. There were no other significant main effects or interactions (all $F < 2.18$, all $p > 0.113$).

Face height

There was a main effect for attribute level, $F(2, 1842) = 26.88$, $p < 0.001$, $\eta^2 = 0.03$. Utility estimates were greater for unaltered ($M = 0.24$, $SD = 1.07$) versus masculinized ($M = -0.10$, $SD = 0.84$, $p < 0.001$, $d = 0.21$) and feminized ($M = -0.14$, $SD = 1.06$, $p < 0.001$, $d = 0.24$) face height. This was moderated by a significant interaction with sex of face, $F(2, 1842) = 68.42$, $p < 0.001$, $\eta^2 = 0.07$. Utility estimates for masculinized face height were greater for male ($M = 0.25$, $SD = 1.65$) than female ($M = -0.45$, $SD = 1.65$) faces, $t(921) = 9.71$, $p < 0.001$, $d = 0.30$. Likewise, estimates for feminized face height were lower for male ($M = -0.43$, $SD = 1.45$) than female ($M = 0.15$, $SD = 1.50$) faces, $t(921) = -8.60$, $p < 0.001$, $d = 0.29$. There were no other significant main effects or interactions (all $F < 1.24$, $p > 0.290$).

Jawbone prominence

There was a main effect for attribute level, $F(2, 1842) = 101.85$, $p < 0.001$, $\eta^2 = 0.10$. Utility estimates were greater for masculinized ($M = 0.26$, $SD = 0.88$, $p < 0.001$, $d = 0.40$) and feminized ($M = 0.17$, $SD = 0.84$, $p < 0.001$, $d = 0.36$) compared to unaltered ($M = -0.43$, $SD = 1.02$) jawbone prominence. This was moderated by a significant interaction between sex of face and attribute level, $F(2, 1842) = 10.34$, $p < 0.001$, $\eta^2 = 0.01$. Utility estimates for masculinized jawbone prominence were greater for male ($M = 0.36$, $SD = 1.15$) than for female ($M = 0.16$, $SD = 1.12$) faces, $t(921) = 4.19$, $p < 0.001$, $d = 0.14$, whereas estimates for feminized jawbone prominence were greater for female ($M = 0.28$, $SD = 1.19$) than for male ($M = 0.06$, $SD = 1.21$) faces, $t(921) = -3.88$, $p < 0.001$, $d = 0.13$. There were no other significant main effects or interactions (all $F < 1.30$, all $p > 0.273$).

Masculinity Ratings

Importance Values

A 2(sex of face [male, female]) \times 5(facial attribute [eyebrow thickness, cheekbone prominence, eye size, face height, jawbone prominence]) repeated-measures ANOVA was used to examine differences in importance values for each facial attribute in male and female faces ranked for masculinity. There was a main effect for facial attribute, $F(4, 3684) = 17.61$, $p < 0.001$, $\eta^2 = 0.02$. Importance values for eyebrow thickness ($M = 21.09$, $SD = 8.65$), face height ($M = 21.13$, $SD = 10.23$), and jawbone prominence ($M = 20.38$, $SD = 7.49$) were not significantly different, but each was significantly greater than cheekbone prominence ($M = 18.34$, $SD = 6.47$; $d = 0.23$; $d = 0.21$; $d = 0.21$, respectively) and eye size ($M = 19.07$, $SD = 6.98$; $d = 0.16$; $d = 0.15$; $d = 0.11$, respectively) (all $ps < 0.001$). This was moderated by a significant interaction with sex of face, $F(4, 3684) = 5.11$, $p < 0.001$, $\eta^2 = 0.01$. Importance values for face height were greater for female ($M = 22.09$, $SD = 13.18$) than for male ($M = 20.17$, $SD = 11.76$) faces, $t(921) = 4.06$, $p < 0.001$, $d = 0.14$. By contrast, importance values for jawbone prominence were greater for male ($M = 20.90$, $SD = 9.74$) than for female ($M = 19.86$, $SD = 10.21$) faces, though this difference was only marginally significant after Bonferroni correction, $t(921) = 2.39$, $p = 0.017$, $d = 0.08$.

Utility Estimates

Five 2(sex of face [male, female]) \times 3(attribute level [masculinized, unaltered, feminized]) repeated-measures ANOVAs were used to examine differences in utility estimates for each level of each facial attribute in male and female faces ranked for perceived masculinity.

Eyebrow thickness

There was a main effect for attribute level, $F(2, 1842) = 77.97$, $p < 0.001$, $\eta^2 = 0.08$. Utility estimates were greater for masculinized ($M = 0.43$, $SD = 1.29$) than unaltered ($M = 0.08$, $SD = 1.26$, $p < 0.001$, $d = 0.17$) and feminized ($M = -0.51$, $SD = 1.46$, $p < 0.001$, $d = 0.39$) eyebrow thickness. There were no other significant main effects or interactions (all $F < 0.88$, all $p > 0.417$).

Cheekbone prominence

There was a main effect for attribute level, $F(2, 1842) = 10.73$, $p < 0.001$, $\eta^2 = 0.01$. Utility estimates were greater for masculinized ($M = 0.17$, $SD = 1.09$) than for feminized ($M = -0.07$, $SD = 1.19$, $p = 0.001$, $d = 0.09$) and unaltered ($M = -0.11$, $SD = 1.10$, $p < 0.001$, $d = 0.11$) cheekbone prominence. There was also a significant interaction between sex of face and attribute level, $F(2, 1842) = 3.17$, $p = 0.042$, $\eta^2 = 0.003$. Utility estimates for feminized cheekbone prominence were greater for female ($M = 0.02$, $SD = 1.64$) than for male ($M = -0.15$, $SD = 1.64$) faces, though this was only marginally significant after Bonferroni correction, $t(921) = 2.34$, $p = 0.019$, $d = 0.08$.

Eye size

There was a main effect for attribute level, $F(2, 1842) = 29.53$, $p < 0.001$, $\eta^2 = 0.03$. Utility estimates were greater for masculinized ($M = 0.26$, $SD = 1.02$) than for unaltered ($M = 0.00$, $SD = 1.23$, $p < 0.001$, $d = 0.15$) and feminized ($M = -0.26$, $SD = 1.32$, $p < 0.001$, $d = 0.25$) eye size. Estimates were also greater for unaltered than for feminized eye size ($p = 0.002$, $d = 0.11$). There were no other significant main effects or interactions (all $F < 0.91$, $p > 0.401$).

Face height

There was a main effect for attribute level, $F(2, 1842) = 147.61$, $p < 0.001$, $\eta^2 = 0.14$. Utility estimates were greater for masculinized ($M = 0.72$, $SD = 1.57$) than for unaltered ($M = -0.05$, $SD = 1.12$, $p < 0.001$, $d = 0.34$) and feminized ($M = -0.66$, $SD = 1.49$, $p < 0.001$, $d = 0.49$) face height. Similarly, estimates were greater for unaltered compared to feminized face height ($p < 0.001$, $d = 0.28$). There was also a significant interaction between sex of face and attribute level, $F(2, 1842) = 3.10$, $p = 0.045$, $\eta^2 = 0.003$. Utility estimates for masculinized face height were greater for female ($M = 0.81$, $SD = 2.14$) than for male ($M = 0.63$, $SD = 1.78$) faces, $t(921) = -2.30$, $p = 0.022$, $d = 0.08$, whereas estimates for unaltered face height were greater for male ($M = 0.02$, $SD = 1.49$) than for female ($M = -0.13$, $SD = 1.55$) faces, $t(921) = 2.09$, $p = 0.037$, $d = 0.07$. However, both of these pairwise comparisons were nonsignificant after Bonferroni correction.

Jawbone prominence

There was a main effect for attribute level, $F(2, 1842) = 44.30$, $p < 0.001$, $\eta^2 = 0.05$. Utility estimates were greater for masculinized ($M = 0.35$, $SD = 1.15$) than for unaltered ($M = -0.31$, $SD = 1.32$, $p < 0.001$, $d = 0.30$) and feminized ($M = -0.04$, $SD = 1.23$, $p < 0.001$, $d = 0.20$) jawbone prominence. Estimates were also higher for feminized than for unaltered jawbone prominence ($p < 0.001$, $d = 0.12$). There were no other significant main effects or interactions (all $F < 1.14$, $p > 0.321$).

DISCUSSION

The relative importance of five facial features (i.e., eyebrow thickness, cheekbone prominence, eye size, face height, and jawbone prominence) to perceptions of physical attractiveness and masculinity were assessed during participants' evaluations of potential romantic partners' facial photographs. CA was used to calculate individual facial feature importance values in overall rankings of attractiveness and masculinity and utility estimates for each attribute. Importance values for perceived masculinity were not significantly different for eyebrow thickness, face height, or jawbone prominence, but each of these traits was significantly greater than cheekbone prominence and eye size, suggesting that perceptions of physical masculinity are more strongly influenced by eyebrow thickness, face height, and jawbone prominence.

compared to cheekbone prominence and eye size. Interestingly, this interacted with the sex of the face being ranked, whereby importance values for face height were greater for female faces compared to male faces. This indicates that a masculinized face height has a greater impact on the perceived masculinity of women's faces than men's faces. The opposite was true of jawbone prominence, which was perceptually more important in attributing masculinity to male faces than female faces. This latter finding should be interpreted with caution, however, as the effect fell short of significance after Bonferroni correction. Finally, utility estimates for masculinity rankings indicated that all features manipulated to appear more masculine or more feminine were ranked as such. Importantly, this indicates that the transformations were perceived as intended, further validating the use of computer graphics methods in objectively manipulating perceptions of sexual dimorphism (see also, e.g., Welling et al., 2007).

For physical attractiveness, these estimates were compared across (1) long- and short-term relationship contexts, and (2) sex of the face. With respect to physical attractiveness importance values, eyebrow thickness was not significantly more important than face height or jawbone prominence, but both eyebrow thickness and jawbone prominence were more important than cheekbone prominence and eye size when ranking faces for attractiveness. Likewise, importance values were greater for face height than for cheekbone prominence. In other words, participants in the current sample weighted the appearance of eyebrow thickness, face height, and jawbone prominence as most important in determining overall attractiveness, and eyebrow thickness and jawbone prominence as more important than eye size. Each of these traits appear to be important during zero-acquaintance assessments of physical attractiveness (Cunningham, 1986; Baudouin and Tiberghien, 2004), social dominance and maturity (Keating, 1985; Cunningham et al., 1990), and personality (Paunonen et al., 1999), however, this was the first study to examine which features are relatively more salient when digitally altered and presented alongside other digitally altered facial features within the same facial identity. Furthermore, no study has isolated the precise informational value of each particular facial feature to overall potential partner assessment. In this sense, it is difficult to conclude precisely why certain features were prioritized over others, although it opens up promising avenues for future investigation. Therefore, the informational value of each trait to perceptions of attractiveness may be best interpreted with respect to how preference for these traits varies across relationship context and sex of the face.

Importance values revealed that eye size was relatively more important for male versus female faces. Specifically, utility estimates showed that unaltered eye size was preferred more than masculinized or feminized eye size, and masculinized eye size was preferred more than feminized eye size, when collapsing across sex of face. However, masculinized (i.e., smaller) eye size was preferred more within short- versus long-term relationship contexts. One possibility is that eye size is an important attractiveness indicator in men insofar as individuals with smaller (i.e., masculinized) eyes are perceived as more mature or more socially dominant (Keating, 1985), which may make them more

desirable as potential sexual (but not necessarily long-term) partners. Interestingly, unaltered eye size was preferred more in a long- compared to short-term context, perhaps suggesting that only a moderate level of masculinity/femininity is preferred in a long-term mate. However, these latter two findings fell short of significance after Bonferroni correction and should be interpreted with caution.

Importance values for face height were greater for female faces than for male faces, suggesting that the overall signal value of face height is more valuable for attributions of female attractiveness than for attributions of male attractiveness. Utility estimates revealed that masculinized face height was considered more attractive for male versus female faces, and, correspondingly, that feminized face height was preferred less for male versus female faces. Thus, sex-typical face height is preferred in both sexes, but the value of this trait in general is more salient in women's faces compared to men's, likely reflecting greater preferences for sex-typical women versus men (see Little et al., 2014). Similarly, utility estimates for masculinized jawbone prominence were greater for male than for female faces, and, correspondingly, estimates for feminized jawbone prominence were greater for female than for male faces, again suggesting that sex-typical traits are preferred more than sex-atypical traits. However, utility estimates were greater for masculinized compared to unaltered and feminized cheekbone prominence, regardless of the sex of the faces being ranked. Although sensible for male faces, this finding is unexpected for female faces, among whom one would expect a higher preference for feminized (i.e., more prominent) cheekbones. It is possible that participants simply preferred more masculine cheekbone prominence in both sexes, but this interpretation should be taken with caution because the highly subtle nature of this particular manipulation may have also made the differences more difficult to discern for the participants. That said, participants did accurately perceive masculinized cheekbone prominence as more masculine than unaltered or feminized cheekbone prominence, suggesting the manipulation was not imperceptibly subtle. This relationship should be investigated more thoroughly in future work.

Finally, although importance values for eyebrow thickness were greater for a long-term compared to short-term relationship context, this relationship was not significant after Bonferroni correction. Surprisingly, utility estimates revealed that thicker eyebrows were more attractive for female faces than male faces, particularly within a long-term mating context. Previous research shows that thicker eyebrows are typically perceived as more masculine and dominant (Windhager et al., 2011) and are more attractive in male than female faces (Keating, 1985), making this finding unexpected. It is possible that unmeasured personality features (e.g., sociosexuality, self-rated attractiveness) of the raters may moderate these findings. Indeed, recent work suggests that eyebrows may signal personality qualities (e.g., narcissism; Giacomini and Rule, 2018) that influence perceptions of attractiveness. Additionally, effect sizes in the current study were relatively small, and so investigation into the mediating effects of individual difference variables may further explain this relationship. Alternatively, it is possible that current and/or

temporary cosmetic and style trends popular among the tested cohort are influencing this relationship, and so this relationship may not generalize to other populations. Certainly, this requires further investigation.

CONCLUSION

This study is the first to demonstrate the utility of CA in investigating the importance of specific aspects or traits of a larger construct (e.g., facial masculinity) to the overall evaluation of that construct. Using this technique, we identified three traits (i.e., eyebrow thickness, jawbone prominence, and face height) whose digital manipulation appears to exert a relatively greater influence on perceptions of romantic partner attractiveness and masculinity than cheekbone prominence and eye size. We also showed how the relative salience of these traits shifts across the sex of the face and the relationship context (i.e., a long-term committed versus purely sexual relationship) for which they are evaluated. This contributes to a burgeoning literature of data-driven face analyses and promises to enrich the development

of methodologies that allow researchers to study the relative contribution of distinct facial features to perceptually holistic facial representations.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Oakland University Institutional Review Board with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Oakland University Institutional Review Board.

AUTHOR CONTRIBUTIONS

JM was responsible for devising the study design, data collection and analysis, and manuscript preparation. LW provided advisory support and guidance throughout each step of this project.

REFERENCES

- Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., and Little, A. C. (2008). Testosterone and financial risk preferences. *Evol. Hum. Behav.* 29, 384–390. doi: 10.1016/j.evolhumbehav.2008.07.001
- Baudouin, J. Y., and Tiberghien, G. (2004). Symmetry, averageness, and feature size in the facial attractiveness of women. *Acta Psychol.* 117, 313–332. doi: 10.1016/j.actpsy.2004.07.002
- Buss, D. M., and Schmitt, D. P. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychol. Rev.* 100:204. doi: 10.1037/0033-295X.100.2.204
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: quasi-experiments on the sociobiology of female facial beauty. *J. Pers. Soc. Psychol.* 50, 925–935. doi: 10.1037/0022-3514.50.5.925
- Cunningham, M. R., Barbee, A. P., and Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *J. Pers. Soc. Psychol.* 59, 61–72. doi: 10.1037/0022-3514.59.1.61
- DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proc. R. Soc. Lond. B Biol. Sci.* 273, 1355–1360. doi: 10.1098/rspb.2005.3445
- Funk, F., Walker, M., and Todorov, A. (2017). Modelling perceptions of criminality and remorse from faces using a data-driven computational approach. *Cogn. Emot.* 31, 1431–1443.
- Gangestad, S. W., and Scheyd, G. J. (2005). The evolution of human physical attractiveness. *Annu. Rev. Anthropol.* 34, 523–548. doi: 10.1146/annurev.anthro.33.070203.143733
- Gangestad, S. W., and Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *Behav. Brain Sci.* 23, 573–587. doi: 10.1017/S0140525X0000337X
- Giacomin, M., and Rule, N. O. (2018). Eyebrows cue grandiose narcissism. *J. Pers.* doi: 10.1111/jopy.12396 [Epub ahead of print].
- Gustafsson, A., Herrmann, A., and Huber, F. (2007). *Conjoint Measurement: Methods and Applications*. Berlin: Springer Science & Business Media. doi: 10.1007/978-3-540-71404-0
- Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate Data Analysis With Readings*. Englewood Cliffs, NJ: Prentice Hall.
- Hu, S., Xiong, J., Fu, P., Qiao, L., Tan, J., Jin, L., et al. (2017). Signatures of personality on dense 3D facial images. *Sci. Rep.* 7:73. doi: 10.1038/s41598-017-00071-5
- Jack, R. E., and Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annu. Rev. Psychol.* 68, 269–297. doi: 10.1146/annurev-psych-010416-044242
- Jones, A. L. (2018). The influence of shape and colour cue classes on facial health perception. *Evol. Hum. Behav.* 39, 19–29. doi: 10.1016/j.evolhumbehav.2017.09.005
- Jones, A. L., Robin, S. S. K., and Ward, R. (2012). Signals of personality and health: the contributions of facial shape, skin texture, and viewing angle. *J. Exp. Psychol.* 38, 1353–1361. doi: 10.1037/a0027078
- Jones, B. C., Little, A. C., Boothroyd, L., DeBruine, L. M., Feinberg, D. R., Smith, M. L., et al. (2005). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Horm. Behav.* 48, 283–290. doi: 10.1016/j.yhbeh.2005.03.010
- Keating, C. F. (1985). Gender and the physiognomy of dominance and attractiveness. *Soc. Psychol. Q.* 48, 61–70. doi: 10.2307/3033782
- Koehler, N., Simmons, L. W., Rhodes, G., and Peters, M. (2004). The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proc. R. Soc. Lond. B Biol. Sci.* 271, S233–S236. doi: 10.1098/rsbl.2003.0146
- Little, A. C., Apicella, C. L., and Marlowe, F. W. (2007). Preferences for symmetry in human faces in two cultures: Data from the UK and the Hadza, an isolated group of hunter-gatherers. *Proc. R. Soc. B Biol. Sci.* 274, 3113–3117. doi: 10.1098/rspb.2007.0895
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011a). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Little, A. C., McPherson, J., Dennington, L., and Jones, B. C. (2011b). Accuracy in assessment of self-reported stress and a measure of health from static facial information. *Pers. Individ. Differ.* 51, 693–698. doi: 10.1016/j.paid.2011.06.010
- Little, A. C., Jones, B. C., Feinberg, D. R., and Perrett, D. I. (2014). Men's strategic preferences for femininity in female faces. *Br. J. Psychol.* 105, 364–381. doi: 10.1111/bjop.12043
- McKone, E., and Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychon. Bull. Rev.* 16, 778–797. doi: 10.3758/PBR.16.5.778
- Mogilski, J. K., Wade, T. J., and Welling, L. L. (2014). Prioritization of potential mates' history of sexual fidelity during a conjoint ranking task. *Pers. Soc. Psychol. Bull.* 40, 884–897. doi: 10.1177/0146167214529798
- Mogilski, J. K., and Welling, L. L. (2017). The relative importance of sexual dimorphism, fluctuating asymmetry, and color cues to health during evaluation

- of potential partners' facial photographs. *Hum. Nat.* 28, 53–75. doi: 10.1007/s12110-016-9277-4
- Paunonen, S. V., Ewan, K., Earthy, J., Lefave, S., and Goldberg, H. (1999). Facial features as personality cues. *J. Pers.* 67, 555–583. doi: 10.1111/1467-6494.00065
- Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., et al. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proc. R. Soc. Lond. B Biol. Sci.* 268, 1617–1623. doi: 10.1098/rspb.2001.1703
- Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., et al. (1999). Menstrual cycle alters face preference. *Nature* 399, 741–742. doi: 10.1038/21557
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., and Edwards, R. (1999). Symmetry and human facial attractiveness. *Evol. Hum. Behav.* 20, 295–307. doi: 10.1016/S1090-5138(99)00014-8
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Petrican, R., Todorov, A., and Grady, C. (2014). Personality at face value: facial appearance predicts self and other personality judgments among strangers and spouses. *J. Nonverbal Behav.* 38, 259–277. doi: 10.1007/s10919-014-0175-3
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Scheib, J. E., Gangestad, S. W., and Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proc. R. Soc. Lond. B Biol. Sci.* 266, 1913–1917. doi: 10.1098/rspb.1999.0866
- Schwaninger, A., Wallraven, C., and Bülthoff, H. H. (2004). Computational modeling of face recognition based on psychophysical experiments. *Swiss J. Psychol.* 63, 207–215. doi: 10.1024/1421-0185.63.3.207
- Scott, I. M., Pound, N., Stephen, I. D., Clark, A. P., and Penton-Voak, I. S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PLoS One* 5:e13585. doi: 10.1371/journal.pone.0013585
- Stephen, I. D., Hiew, V., Coetsee, V., Tiddeman, B. P., and Perrett, D. I. (2017). Facial shape analysis identifies valid cues to aspects of physiological health in Caucasian, Asian, and African populations. *Front. Psychol.* 8:1883. doi: 10.3389/fpsyg.2017.01883
- Stephen, I. D., Scott, I. M., Coetsee, V., Pound, N., Perrett, D. I., and Penton-Voak, I. S. (2012). Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol. Hum. Behav.* 33, 260–267. doi: 10.1016/j.evolhumbehav.2011.10.003
- Stephen, I. D., Smith, M. J. L., Stirrat, M. R., and Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *Int. J. Primatol.* 30, 845–857. doi: 10.1007/s10764-009-9380-z
- Thornhill, R., and Gangestad, S. W. (1999). Facial attractiveness. *Trends Cogn. Sci.* 3, 452–460. doi: 10.1016/S1364-6613(99)01403-5
- Thornhill, R., and Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evol. Hum. Behav.* 27, 131–144. doi: 10.1016/j.evolhumbehav.2005.06.001
- Todorov, A. (2017). *Face Value: The Irresistible Influence of First Impressions*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400885725
- Todorov, A., Dotsch, R., Wigboldus, D. H., and Said, C. P. (2011). Data-driven methods for modeling social perception. *Soc. Pers. Psychol. Compass* 5, 775–791. doi: 10.1111/j.1751-9004.2011.00389.x
- Todorov, A., Olivola, C. Y., Dotsch, R., and Mende-Siedlecki, P. (2015). Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu. Rev. Psychol.* 66, 519–545. doi: 10.1146/annurev-psych-113011-143831
- Todorov, A., Said, C. P., Engell, A. D., and Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* 12, 455–460. doi: 10.1016/j.tics.2008.10.001
- Walker, M., and Vetter, T. (2016). Changing the personality of a face: perceived big two and big five personality factors modeled in real photographs. *J. Pers. Soc. Psychol.* 110, 609–624. doi: 10.1037/pspp0000064
- Welling, L. L. M., Jones, B. C., and DeBruine, L. M. (2008). Sex drive is positively associated with women's preferences for sexual dimorphism in men's and women's faces. *Pers. Individ. Differ.* 44, 161–170. doi: 10.1016/j.paid.2007.07.026
- Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Smith, M. L., Little, A. C., et al. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Horm. Behav.* 52, 156–161. doi: 10.1016/j.yhbeh.2007.01.010
- Windhager, S., Schaefer, K., and Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *Am. J. Hum. Biol.* 23, 805–814. doi: 10.1002/ajhb.21219
- Zebrowitz, L. A. (2011). "Ecological and social approaches to face perception," in *The Oxford Handbook of Face Perception*, eds A. Calder, G. Rhodes, M. Johnson, and J. Haxby (Oxford: Oxford University Press), 31–50.
- Zhang, H., Li, Q., Sun, Z., and Liu, Y. (2018). Combining data-driven and model-driven methods for robust facial landmark detection. *IEEE Trans. Inform. Forensics Secur.* 13, 2409–2422. doi: 10.1109/TIFS.2018.2800901

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Mogilski and Welling. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



360 Degrees of Facial Perception: Congruence in Perception of Frontal Portrait, Profile, and Rotation Photographs

Vít Třebický^{1,2*}, Jitka Fialová^{1,2}, David Stella^{1,2}, Zuzana Štěrbová^{1,2}, Karel Kleisner^{1,2} and Jan Havlíček^{1,2}

¹ National Institute of Mental Health, Klecany, Czechia, ² Faculty of Science, Charles University, Prague, Czechia

OPEN ACCESS

Edited by:

Ian Stephen,
Macquarie University, Australia

Reviewed by:

Barnaby James Wyld Dixon,
The University of Queensland,
Australia

Danielle Leigh Wagstaff,
Federation University, Australia

*Correspondence:

Vít Třebický
vit.trebicky@natur.cuni.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 September 2018

Accepted: 15 November 2018

Published: 07 December 2018

Citation:

Třebický V, Fialová J, Stella D,
Štěrbová Z, Kleisner K and Havlíček J
(2018) 360 Degrees of Facial
Perception: Congruence in Perception
of Frontal Portrait, Profile, and
Rotation Photographs.
Front. Psychol. 9:2405.
doi: 10.3389/fpsyg.2018.02405

Studies in social perception traditionally use as stimuli frontal portrait photographs. It turns out, however, that 2D frontal depiction may not fully capture the entire morphological diversity of facial features. Recently, 3D images started to become increasingly popular, but whether their perception differs from the perception of 2D has not been systematically studied as yet. Here we investigated congruence in the perception of portrait, left profile, and 360° rotation photographs. The photographs were obtained from 45 male athletes under standardized conditions. In two separate studies, each set of images was rated for formidability (portraits by 62, profiles by 60, and 360° rotations by 94 raters) and attractiveness (portraits by 195, profiles by 176, and 360° rotations by 150 raters) on a 7-point scale. The ratings of the stimuli types were highly intercorrelated (for formidability all $r_s > 0.8$, for attractiveness all $r_s > 0.7$). Moreover, we found no differences in the mean ratings between the three types of stimuli, neither in formidability, nor in attractiveness. Overall, our results clearly suggest that different facial views convey highly overlapping information about structural facial elements of an individual. They lead to congruent assessments of formidability and attractiveness, and a single angle view seems sufficient for face perception research.

Keywords: 2D, 3D, head, standardized photography, assessment, morphology, attractiveness, formidability

INTRODUCTION

When artists create portraits, they rarely depict a full frontal view of the face of a given sitter. Instead, they tend to portray people in some degree of profile, emphasizing one cheek and dimensionality of a face (Murphy, 1994). Interestingly, vast majority of studies on facial perception uses frontal portraits (un/altered photographs, morphs, or line drawings) as stimuli (e.g., Thornhill and Gangestad, 1999; Rhodes, 2006; Kościński, 2009; Calder et al., 2011; Little et al., 2011; Valentová et al., 2013; Little, 2014). Given, however, that in our daily lives we experience faces from multiple angles, it is far from certain that a frontal view is the optimal depiction and several studies even suggested that an individual's appearance can significantly vary depending on the viewing angle (Rule et al., 2009; Jenkins et al., 2011; Tigue et al., 2012; Kościński and Zalewska, 2017; Sutherland et al., 2017). Faces are, after all, complex and highly variable morphological structures (Enlow et al., 1996) and some facial features are apparent only from some viewing angles (Danel et al., 2018). For example, Danel et al. (2018) reported only a moderate correlation in sexually dimorphic features between lateral and frontal facial configuration in both men and women. When frontal and lateral facial configurations were compared as to their averageness,

a significant association was found only in women. It is therefore plausible to assume that complementary information may be provided by different viewing angles. A single frontal view could potentially obscure relevant visual cues used in assessing certain dimensions (e.g., determinants of facial masculinity, such as protrusion of the brow ridge and angularity of the jaw), thus reducing judgment accuracy.

In research on body perception, the use of other than just frontal view is becoming increasingly common (Tovée and Cornelissen, 2001; Perilloux et al., 2012; Sell et al., 2017; Cornelissen et al., 2018). The use of multiple body angles views allows for assessments of multivariate trait interactions (Brooks et al., 2015). Varying viewing angles of bodies allow raters to assess the shapes and sizes of various morphological characteristics, such as body fat, lean mass distribution, or breast morphology (Dixson et al., 2011, 2015) which all contribute to the resulting attractiveness rating.

Research on facial perception that employs other than frontal facial views remains, however, at best unsystematic (Kościński, 2009) and mutual relations between the frontal and lateral dimensions of facial features have so far received very little attention (Danel et al., 2018). Profile views have been used primarily in orthodontics and aesthetic medicine because it is known that they have an impact on facial attractiveness judgments (Spyropoulos and Halazonetis, 2001; Johnston et al., 2005; Maple et al., 2005; Soh et al., 2007; Shafiee et al., 2008; Nomura et al., 2009). Results from several studies that investigated the averageness of facial profiles show patterns analogous to frontal images (Spyropoulos and Halazonetis, 2001; Minear and Park, 2004; Valentine et al., 2004; Valenzano et al., 2006). Some researchers, meanwhile, tried to overcome the limitations of a single view stimulus by presenting raters with both frontal and lateral views of targets on a single screen (e.g., Dixson and Rantala, 2015; Dixson et al., 2016; Valentova et al., 2017), while other studies found a medium to high correlation between the rating of attractiveness of frontal and lateral depictions (ranging from $r = 0.52$ to 0.83) (Diener et al., 1995; Valenzano et al., 2006; Davidenko, 2007; Shafiee et al., 2008; Kościński and Zalewska, 2017).

Until recently, most studies used as stimuli static, two-dimensional images (photographs). Thanks to technological progress, including a considerable increase in computers' computing powers, 3D scanning and 3D reconstruction technology is now becoming more accessible to facial perception research (Toole et al., 1999; Caharel et al., 2009; Chelnokova and Laeng, 2011; Meyer-Marcotty et al., 2011; Jones et al., 2012; Lefevre et al., 2012; Tigue et al., 2012; Berssenbrügge et al., 2014; Mydlová et al., 2015; Holzleitner and Perrett, 2016; Hu et al., 2017; Kordsmeyer et al., 2018). Potential bias associated with a single 2D image (e.g., profile) might be minimized by the use of 3D images, which represent various viewing angles. To our best knowledge, however, only one study directly compared ratings based on 2D and 3D facial images (Tigue et al., 2012). Authors found a high correlation between 2D and 3D stimuli ($r = 0.71$), with mean ratings significantly higher for 3D images. On the other hand, it should be noted that in this study, only opposite-sex ratings were performed (female faces were rated by

male participants), on a single scale (attractiveness), and the only 2D depictions used were frontal portraits.

Current evidence suggests a rather high level of congruence in judged characteristics (especially attractiveness) between frontal and lateral or frontal and 3D views of faces. It should, however, be taken into account that the development of morphological features between the frontal and lateral view does not always correlate (Danel et al., 2018) and one could thus expect that some socially relevant traits may be easier to assess from other than frontal view (Tigue et al., 2012).

In the two studies, we estimated the congruence in perception of three different views of male heads (frontal portrait, left profile, and 360° rotation photographs). We employed two characteristics relevant in the context of intra- and inter-sexual selection, namely the rating of formidability and attractiveness. We also explored whether the type of device used (mobile phones, laptop, and desktop computers) influences the ratings.

MATERIALS AND METHODS

All procedures employed in this study conform to the ethical standards of the relevant committee on human experimentation and with the Helsinki Declaration. The study was approved by the Institutional Review Board of National Institute of Mental Health, Czech Republic (Ref. num. 28/15). All participants were informed about the goals of the study and gave their informed consent. The present study is part of a larger project which investigates multimodal perception of traits associated with sexual selection and characteristics related to competition outcome.

Targets

We collected photographs of 45 male Mixed Martial Arts (MMA) athletes (mean age = 26.6, $SD = 5.86$, range = 18–38). All athletes were from the Czech Republic. They were invited via social media advertisements, leaflets distributed at domestic MMA tournaments, gyms, and with the assistance of Mixed Martial Arts Association Czech Republic (MMAA). All targets were provided with brief description of the project and approved their participation by signing informed consent. As compensation for their participation, they received 400 CZK (approx. €15).

Acquisition and Settings of the Photographs

To capture images of the targets' head from all 360°, we built a turning plywood platform (120 cm in diameter) using flat ball bearings. The platform had 36 steps around its perimeter, i.e., one step for every 10°, making it basically a large turntable. To achieve standardization—all photographs were acquired on site—the platform was placed inside a purpose-built portable photographic booth to control for changes in ambient illumination and for color reflections (see e.g., Rowland and Burriss, 2017; Thorstenson, 2018). Booth dimensions were $140 \times 140 \times 255$ cm. Its frame was made of sectioned aluminum profiles. The outside of the booth and the inside of its roof was covered with black duvetyne cloth (a dense fabric), while the internal side of the walls, the seamless backdrop, and surface

of the turning platform were covered with a bright white velvet (medium density fabric).

To achieve standardized lighting conditions, we used one 800 W studio strobe (Photon Europe MSN HSS-800) aimed into a white reflective umbrella used as a light modifier (Photon Europe, 109 cm diameter), mounted on a 175 cm high light stand, and tilted 10° downwards toward the booth. The light was positioned 125 cm from the target. This lighting setup ensured even exposure across the whole scene, which was further verified before each session by a digital light meter (Sekonic L-308S).

Images were acquired using a 24-megapixel, full-frame (35.9×24 mm CMOS sensor, a 35 mm film equivalent) digital SLR camera Nikon D610 equipped with a fixed focal length lens Nikon AF-S NIKKOR 85 mm f/1.8 G (Třebický et al., 2016). Exposure values were set to ISO 100, shutter speed 1/200 s, and aperture f/11. Photographs were shot into 14-bit uncompressed raw files (NEF) and AdobeRGB color space. Color calibration was performed using X-Rite Color Checker Passport color targets and white balance patch photographed at the beginning of each session. The camera was mounted in portrait orientation directly onto the light stand, which carried also the strobe light positioned 125 cm from the target so as to achieve a perception close to social interpersonal distance (Hall, 1966; Baldassare and Feller, 1975; Sorokowska et al., 2017), to maintain a constant perspective distortion (Třebický et al., 2016; Erkelens, 2018), and to avoid potential perception bias based on interpersonal distance (Bryan et al., 2012). Camera's distance from each target was verified with a digital laser rangefinder (Bosch PLR 15) as distance between the sensor plane (marked ϕ on camera body) to the middle of target's forehead. Camera's height was adjusted for each target so as to position the center of his head in the middle of the frame. Focus point was set on target's right eye and focus distance was locked for further images of the target. This setting of camera's distance, focal length, and sensor size yielded a 35×53 cm field of view (23.85° angle of view) and the aperture setting resulted in a 9 cm depth of field (4 cm before and 5 cm behind the focal plane).

Targets were seated on a 63 cm high bar stool (Ikea Franklin) positioned in the middle (rotation axis) of the turning platform. We asked them not to lean against the stool's back support and to sit with their back straight and hands hanging freely alongside their body. They were asked to adopt a neutral facial expression (with no smile or frown), to look directly into the camera, and to remain in this position for all subsequent photographs. When necessary, targets were instructed to adjust their posture and head position, so they were facing the camera straight on, without any head pitch, yaw, or roll. On top of that, they were instructed to wear only black underwear shorts we provided them with (i.e., without T-shirt) and to remove any adornments (glasses, earrings, piercings or other jewelry).

One full 360° rotation yielded 36 photographs. After each photograph, research assistant manually turned the turning platform by one step (10°) clockwise. We captured two full rotations to obtain one full set of images while eliminating all possible movements, blinks, etc. of targets. Capturing both full rotations took approx. 10 min.

Stimuli Processing

All image processing was carried out in Adobe Lightroom Classic CC (Version 2017). First, all images were converted into DNG raw files and DNG color calibration profiles were assembled (using X-Rite Color Checker Passport LR plugin) and applied to all photographs. For each target, a final set of 36 images covering full 360° head rotation was selected and postprocessed by combining suitable images (correct head position, open eyes, closed mouth, etc.) from the two captured rotations. To ensure consistency in exposure across all selected photographs, percentages of Red, Green, and Blue channel values were checked across three background areas (above, left, right) and eventual small differences in exposure were manually adjusted to the same level. In the next step, the calibrated images were exported into lossless 16-bit AdobeRGB TIFF files in their real size of 35×53 cm and 168 pixels per inch (ppi) resolution (a native ppi of 4K screens used for rating sessions, see Rating Session in section Formidability Rating). This resulted in life-sized images of targets' heads. Horizontal and vertical positions of images were adjusted using LR Transform tool to position target's head in the center of the frame with eyes in a horizontal line. Final images were batch-cropped to 1:1.1 ($2,095 \times 2,305$) side ratio to fit head rotations of all targets. Images were then converted into sRGB color space and exported as 8-bit JPEG files ($2,095 \times 2,305$ px @ 168 ppi).

Building 360° Head Rotations

We used Sirv (www.sirv.com), an online suite for creating and managing image spins, to build 360° head rotations. With all image adjustments and optimization to image size and quality done by Sirv turned off, we uploaded the images of all targets and created the individual spins. See **Supplementary Materials** for sample 360° head rotation (360 rotation video.MP4).

Portraits and Profiles

Analogously to previous research investigating morphological differences between portraits and profiles (Danel et al., 2018), we have selected from the set of 36 images for each target a frontal and left profile image. See **Supplementary Materials** for sample frontal portrait (Frontal portrait.JPEG) and left profile (Left profile.JPEG).

Raters

Formidability Rating

Portraits were evaluated by 62 raters (30 men), mean age = 23.1 ($SD = 3.45$, range = 18–39); profiles by 60 raters (30 men), mean age = 22.8 ($SD = 3.55$, range = 18–36); and 360° rotations by 94 raters (46 men), mean age = 22.1 ($SD = 3.09$, range = 18–38) (see **Table 2**). Raters were mainly Charles University (Prague, Czech Republic) students recruited via social media advertisements, mailing list of participants assembled in previous studies or invited on site. All raters were provided with brief description of the project and approved their participation by signing informed consent. Rating took place in a lab (see section Rating Sessions, Formidability Rating) and when the rating was completed, they received for their participation 100 CZK (approx. €4) and a debriefing leaflet. Using a two-way ANOVA, we found no

age differences between sexes, stimuli type ratings, or their interaction [Sex: $F_{(1,210)} = 0.371$, $p = 0.543$; Stimuli type: $F_{(2,210)} = 1.777$, $p = 0.172$; Sex \times Stimuli type: $F_{(2,210)} = 0.006$, $p = 0.994$].

Attractiveness Rating

Portraits were evaluated by 195 raters (30 men), mean age = 29.6 ($SD = 6.05$, range = 18–48); profiles by 176 raters (32 men), mean age = 29.2 ($SD = 6.26$, range = 18–53); and 360° rotations by 150 raters (35 men), mean age = 29 ($SD = 6.27$, range = 18–46) (see **Table 2**). Raters were recruited mainly via advertisements among followers of National Institute of Mental Health (facebook.com/nudzcz) and Human Ethology group (facebook.com/etologie) Facebook pages. Ratings were carried out online. All raters provided their informed consent by clicking on the “I agree” button to consent with their participation in the study and were not financially reimbursed. Two-way ANOVA showed no age difference between sexes, the stimuli type ratings or their interaction [Sex: $F_{(1,515)} = 2.553$, $p = 0.111$; Stimuli type: $F_{(2,515)} = 0.162$, $p = 0.85$; Sex \times Stimuli type: $F_{(2,515)} = 0.084$, $p = 0.864$]. **Table 2** provides detailed descriptive statistics.

Rating Sessions

Formidability Rating

Formidability ratings were performed in two separate sessions. In the first session, we collected the ratings of 360° rotations. In the second session, raters were randomly divided to rate either a set of portrait or profile images. Each rater thus judged a full set of only one type of stimuli.

Ratings took place in a quiet perception lab, in standardized conditions across all raters (with artificial lighting and closed window blinds to eliminate changes in ambient lighting). Raters were seated in the same eye level with stimuli's eyes, 125 cm from the screen, i.e., in the same distance as the camera was from the target in order to simulate approximate social interpersonal distance (Hall, 1966; Baldassare and Feller, 1975), and in the center of the projected photograph (Cooper et al., 2012). This was implemented so as to increase the ecological validity of the rating.

Images were presented to raters on 27" Dell U2718Q UltraSharp IPS screens ($3,840 \times 2,160$, 99% sRGB color space coverage) turned to a vertical position to accommodate life-size images. Screens were connected to Asus ROG G20 PC running Microsoft Windows 10 with environment scaling set to 100%. Screens were color- and luminance-calibrated with X-Rite i1 Display Pro probes. The probes were connected during the whole rating session to adjust screens for ambient light. Qualtrics survey suite (Qualtrics, Provo, UT) with Blank theme run through Google Chrome (in full screen mode and 100% scaling) was used for data collection.

All raters received a set of brief demographics questions (e.g., sex, age, and education status) followed by a block containing stimuli. Images were presented in a randomized order. Raters were asked to rate formidability (“Jak moc by byl tento muž úspěšný, kdyby se dostal do fyzického souboje?”/“If this man was involved in physical confrontation, how successful he would be?”) of each target on a 7-point verbally anchored

scale (from “1 – velice neúspěšný”/“very unsuccessful,” to “7 – velice úspěšný”/“very successful”). The 360° rotations spun automatically clockwise once (automatic rotation took approx. 2 s) and raters were instructed to turn the heads around for further inspection by dragging mouse left or right before rating. Portrait and profile photographs were simply projected on the screen. Time for rating was not limited.

Attractiveness Rating

Ratings were collected on-line via Qualtrics survey suite (Qualtrics, Provo, UT). All raters were first presented with a brief study description and informed consent. Then they completed a set of demographics questions, which was followed by one randomly selected block of stimuli (portraits, profiles, or 360° rotations). Each rater thus assessed a full set of only one type of stimuli. Images were presented in a randomized order. Raters were asked to rate attractiveness (“Jak atraktivní je muž na fotografii?”/“How attractive is the man on photograph?”) of each target on a 7-point verbally anchored scale (from “1 – velice neatraktivní”/“very unattractive,” to “7 – velice atraktivní”/“very attractive”). The 360° rotations spun automatically once clockwise (automatic rotation took approx. 2 s) and raters were instructed to turn the heads around for further inspection by dragging mouse left or right before rating, while portrait and profile photographs were simply projected on the screen. Time for rating was not limited.

We used Qualtrics *Blank theme* and custom CSS code to set the image size to 800 px width with centered margin alignments (`.Skin #SkinContent.QuestionBody {width: 800px; display: block; margin-left: auto; margin-right: auto;}.Skin #SkinContent.QuestionText {width: 800px; display: block margin-left: auto; margin-right: auto;}`) to standardize stimulus size and position across all devices used. First, raters were asked to switch their web browsers into Full Screen mode and adjust page scaling to achieve the largest image size possible while seeing the rating scale without having to scroll down, i.e., if a Full HD 16:9 screen ($1,920 \times 1,080$) was used for rating in Full Screen mode, browser scaling would remain on native 100%.

Devices used for attractiveness rating

When raters completed rating the images, they were asked to specify the type of device they used (mobile phone, tablet device, laptop computer, desktop computer or other), screen size or brand and model name of the device (to

TABLE 1 | Correlations between stimuli types.

Scale	Stimuli	Pearson's r [95 % CI]
Formidability	Portrait – Profile	0.829 [0.708, 0.903]
	Portrait – 360° rotation	0.974 [0.952, 0.985]
	Profile – 360° rotation	0.882 [0.794, 0.934]
Attractiveness	Portrait – Profile	0.706 [0.520, 0.828]
	Portrait – 360° rotation	0.956 [0.921, 0.976]
	Profile – 360° rotation	0.782 [0.634, 0.875]

All correlations are significant at $p < 0.001$.

later identify screen size and resolution). This data was used to test a possible effect of the device used on the rating.

In total, attractiveness was rated by 521 raters: 233 used laptop computers, 135 desktop computers, 116 mobile phones, 19 tablet devices, 1 other device, and 17 did not specify device type. See **Tables S1** and **S2** for data on screen sizes and resolutions.

In subsequent analyses, we used only data from the three most frequently represented device categories: mobile phones, laptops, and desktop computers. This resulted in a sample of 484 raters.

Statistical Analysis

All statistical tests were performed in JASP 0.9.0.1 (JASP Team, 2018) and jamovi 0.9.1.7 (jamovi project, 2018). McDonald's ω statistics was used for estimating inter-rater agreement (Dunn et al., 2014). To test for potential age differences between rater groups, a two-way ANOVA was carried out, with raters' sex and stimuli types entered as two independent variables and age as a dependent variable. Two-way ANOVA was further used to compare sex differences in mean formidability and attractiveness rating, where raters' sex and stimuli types were entered as two independent variables and the rating of

formidability or attractiveness as a dependent variable. Effect sizes for two-way ANOVAs are reported in η^2 . Associations between the ratings of different stimuli types were tested by bivariate correlations using Pearson's r coefficient with 95% CIs [lower limit, upper limit]. For exploratory purposes, we also tested the influence of device on attractiveness rating using a two-way ANOVA with stimulus type and device type entered as independent variables and mean attractiveness rating as a dependent variable. A Holm's *post-hoc* test was performed and effect sizes for the comparison are reported in Cohen's d .

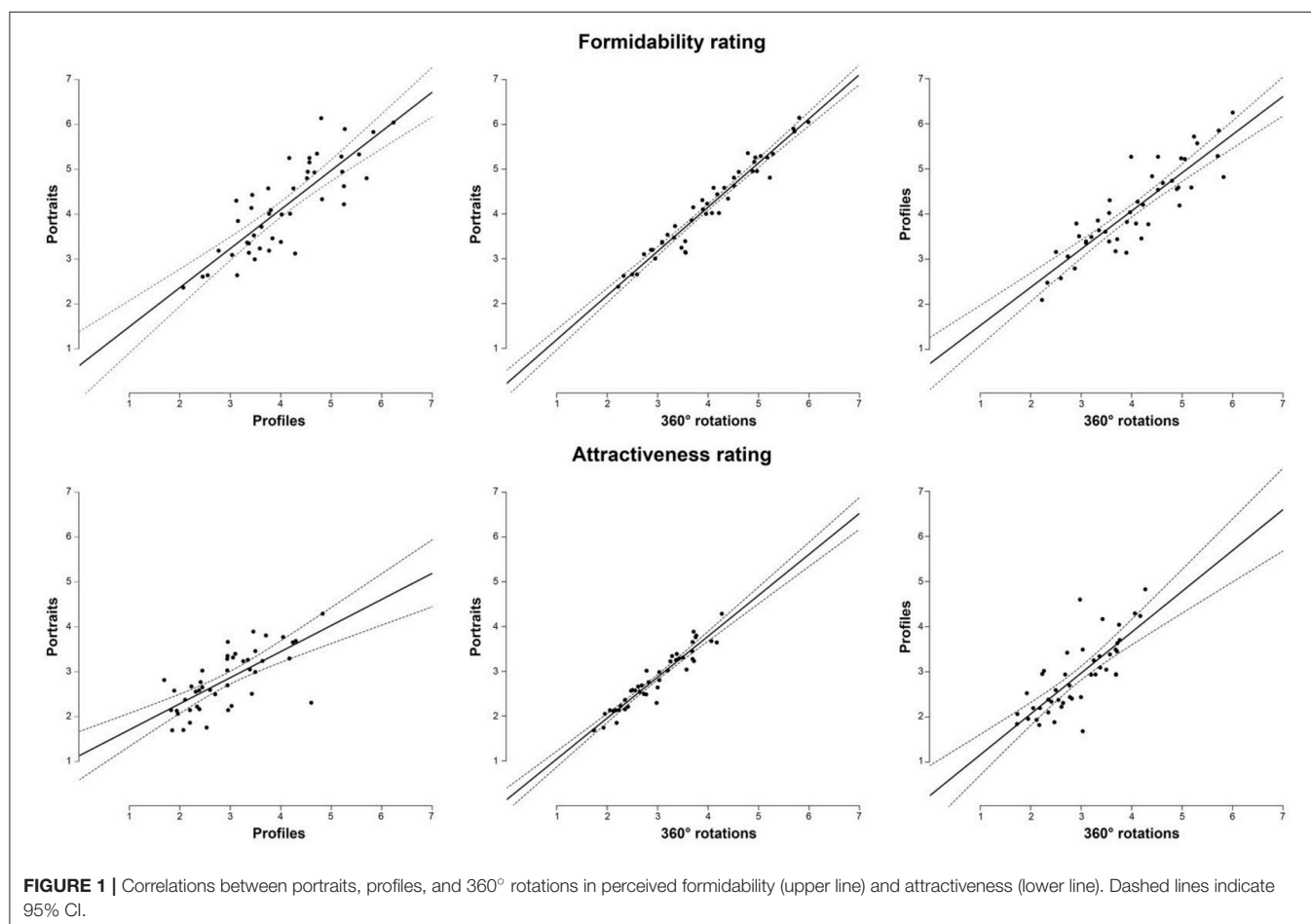
Data Availability

Datasets generated and analyzed during the current study are available as **Supplementary Material** of this article (Dataset formidability.XLSX, Dataset attractiveness.XLSX).

RESULTS

Formidability Rating

McDonald's ω scores of male and female ratings showed a high inter-rater agreement across all three stimuli types (ranging from 0.732 to 0.876). In subsequent analyses, we



have therefore used mean formidability ratings given to the individual stimuli separately by male and female raters. Further, we found a high correlation between ratings assigned by men and women for portraits ($r = 0.941$, 95% CI [0.895, 0.967], $p < 0.001$), profiles ($r = 0.962$, 95% CI [0.931, 0.979], $p < 0.001$) and 360° rotations ($r = 0.972$, 95% CI [0.95, 0.985], $p < 0.001$).

Ratings of all three stimuli types were highly correlated (Table 1, Figure 1). Two-way ANOVA showed no main effect of rater's sex [$F_{(1,264)} = 0.00014$, $p = 0.991$, $\eta^2 < 0.001$], stimulus type [$F_{(2,264)} = 0.473$, $p = 0.624$, $\eta^2 = 0.004$], or rater's sex \times stimulus type interaction [$F_{(2,264)} = 0.01$, $p = 0.99$, $\eta^2 < 0.001$] on formidability ratings (Figure 2). For descriptive statistics, see Table 2.

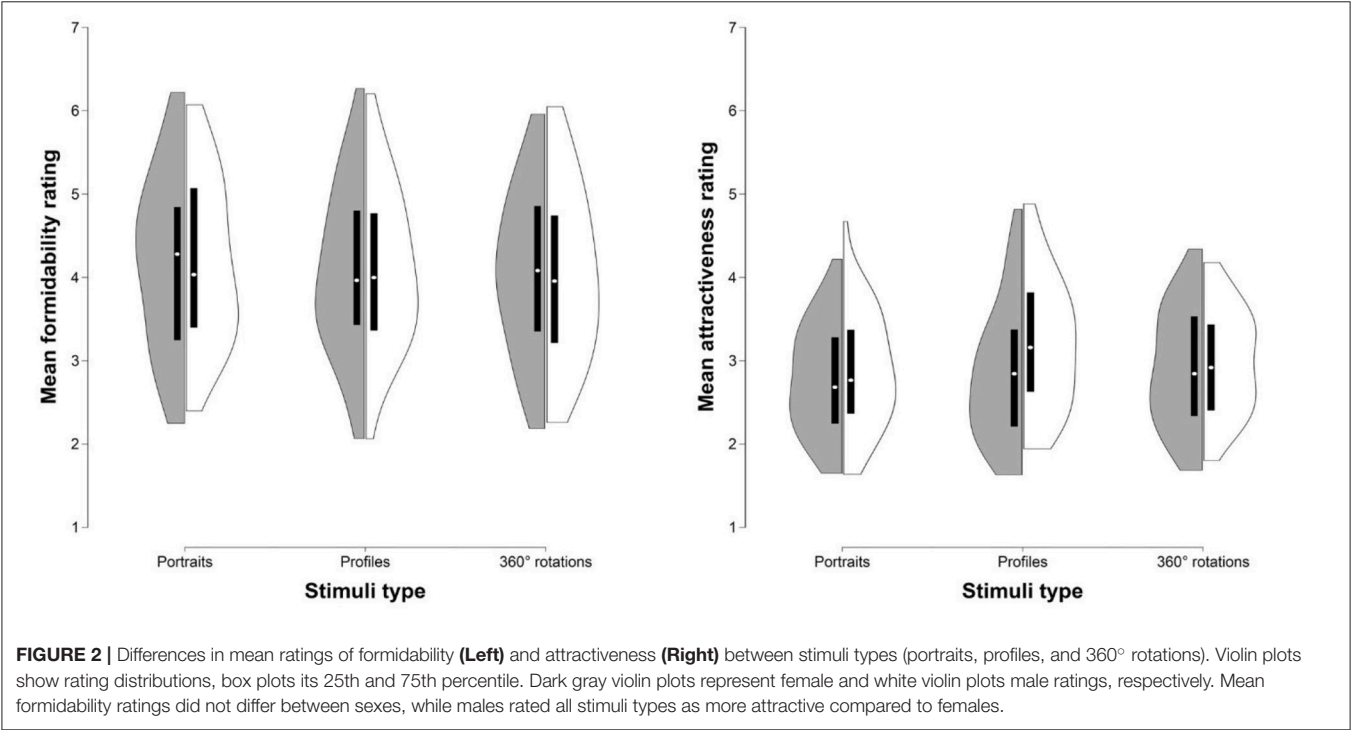


FIGURE 2 | Differences in mean ratings of formidability (Left) and attractiveness (Right) between stimuli types (portraits, profiles, and 360° rotations). Violin plots show rating distributions, box plots its 25th and 75th percentile. Dark gray violin plots represent female and white violin plots male ratings, respectively. Mean formidability ratings did not differ between sexes, while males rated all stimuli types as more attractive compared to females.

TABLE 2 | Descriptive statistics.

Scale	Stimuli	Sex	N	Age			Rating			McDonald's ω
				Mean	SD	Range	Mean	SD	Range	
Formidability	Portraits	Total	62	23.113	3.448	18–39	4.186	1.009	2.25–6.22	0.815
		Female	32	23.219	3.933	18–39	4.193	1.068	2.25–6.22	0.847
		Male	30	23	3.006	20–31	4.179	0.977	2.4–6.07	0.732
	Profiles	Total	60	22.816	3.553	18–36	4.085	0.961	2.07–6.27	0.856
		Female	30	22.967	3.222	19–32	4.073	0.994	2.07–6.27	0.876
		Male	30	22.667	3.907	18–36	4.098	0.946	2.07–6.2	0.844
	360° rotations	Total	94	22.127	3.094	18–38	4.046	0.998	2.19–6.04	0.798
		Female	48	21.854	4.699	18–38	4.049	0.988	2.19–5.96	0.736
		Male	46	21.957	2.556	19–29	4.042	1.022	2.26–6.04	0.857
Attractiveness	Portraits	Total	195	29.6	6.048	18–48	2.803	0.656	1.63–4.67	0.892
		Females	165	29.491	5.854	18–48	2.79	0.654	1.65–4.22	0.9
		Males	30	30.2	7.107	18–41	2.876	0.696	1.63–4.67	0.831
	Profiles	Total	176	29.188	6.257	18–53	2.904	0.796	1.63–4.87	0.962
		Female	144	28.951	6.1	18–46	2.84	0.803	1.63–4.82	0.96
		Male	32	30.25	6.924	18–53	3.194	0.784	1.94–4.87	0.964
	360° rotations	Total	150	29	6.271	18–46	2.926	0.688	1.69–4.34	0.957
		Female	115	28.687	6.353	18–46	2.909	0.715	1.69–4.34	0.954
		Male	35	30.029	5.968	20–43	2.981	0.616	1.8–4.17	0.966

Attractiveness Rating

McDonald's ω scores of male and female ratings showed a high inter-rater agreement in all three stimuli types (ranging from 0.831 to 0.966), which is why in subsequent analyses, we used the mean formidability ratings given to a particular stimulus separately by male and female raters. Ratings by women and men were highly correlated: $r = 0.952$, 95% CI [0.915, 0.974], $p < 0.001$; $r = 0.969$, 95% CI [0.944, 0.983], $p < 0.001$; $r = 0.962$, 95% CI [0.932, 0.979], $p < 0.001$ for portraits, profiles, and 360° rotations, respectively.

Attractiveness ratings of all three stimuli types were highly correlated (Table 1, Figure 1). Two-way ANOVA showed main effect of rater's sex [$F_{(1,264)} = 3.87$, $p = 0.05$, $\eta^2 = 0.014$], men rated attractiveness higher as compared to women, but the effect of stimulus type [$F_{(2,264)} = 1.516$, $p = 0.222$, $\eta^2 = 0.011$], and rater's sex \times stimuli interaction [$F_{(2,264)} = 1.118$, $p = 0.329$, $\eta^2 = 0.008$] on attractiveness ratings was not significant (Figure 2). For descriptive statistics, see Table 2.

Influence of Device Type on Attractiveness Rating

To explore whether the type of device used for viewing and rating influences attractiveness rating, we performed a two-way ANOVA with stimuli type and device type as independent factors. The results showed main effects of both device types [$F_{(2,475)} = 7.429$, $p < 0.001$, $\eta^2 = 0.030$] and stimuli types [$F_{(2,475)} = 4.27$, $p = 0.015$, $\eta^2 = 0.017$], but no significant interaction between them [$F_{(4,475)} = 1.065$, $p = 0.373$, $\eta^2 = 0.008$]. Holm's *post-hoc* comparison showed that raters using mobile phones rated the images as significantly more attractive compared to desktop [$t_{(475)} = 3.817$, $p_{\text{Holm}} < 0.001$, Cohen's $d = 0.557$] and laptop users [$t_{(475)} = 3.023$, $p_{\text{Holm}} = 0.005$, Cohen's $d = 0.392$], whereby the ratings assigned by laptop and desktop users did not differ [$t_{(475)} = 1.357$, $p_{\text{Holm}} = 0.175$, Cohen's $d = 0.145$]. 360° rotations were rated significantly higher than portraits [$t_{(475)} = 2.912$, $p_{\text{Holm}} = 0.011$, Cohen's $d = 0.418$], but there was no statistical difference between 360° rotations and profiles [$t_{(475)} = 1.753$, $p_{\text{Holm}} = 0.16$, Cohen's $d = 0.212$]; and between portraits and profiles [$t_{(475)} = 1.366$, $p_{\text{Holm}} = 0.173$, Cohen's $d = 0.159$] (Figure 3). For descriptive statistics, see Table 3. Further, attractiveness ratings between all three types of devices were highly correlated: $r = 0.883$, 95% CI [0.839, 0.915], $p < 0.001$; $r = 0.885$, 95% CI [0.842, 0.917], $p < 0.001$; $r = 0.949$, 95% CI [0.93, 0.964], $p < 0.001$ for mobile phones–laptops, mobile phones–desktops, and laptops–desktops, respectively (Figure 4).

DISCUSSION

The main aim of this study was to examine whether perception of formidability and attractiveness varies depending on the angle under which a face is viewed. To this purpose, we used standardized sets of frontal portraits, left profiles, and 360° head rotations of male facial images. We found strong correlations between the three types of stimuli and no significant differences

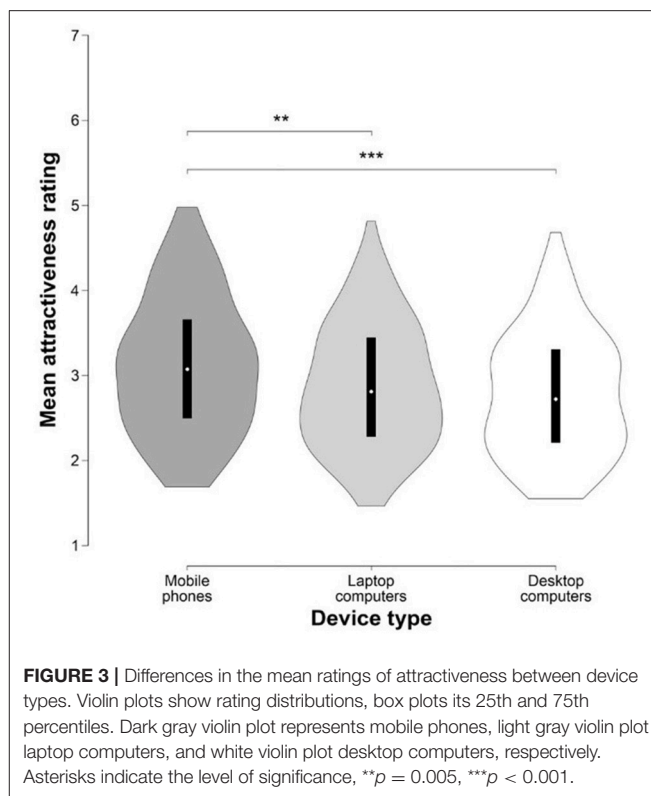
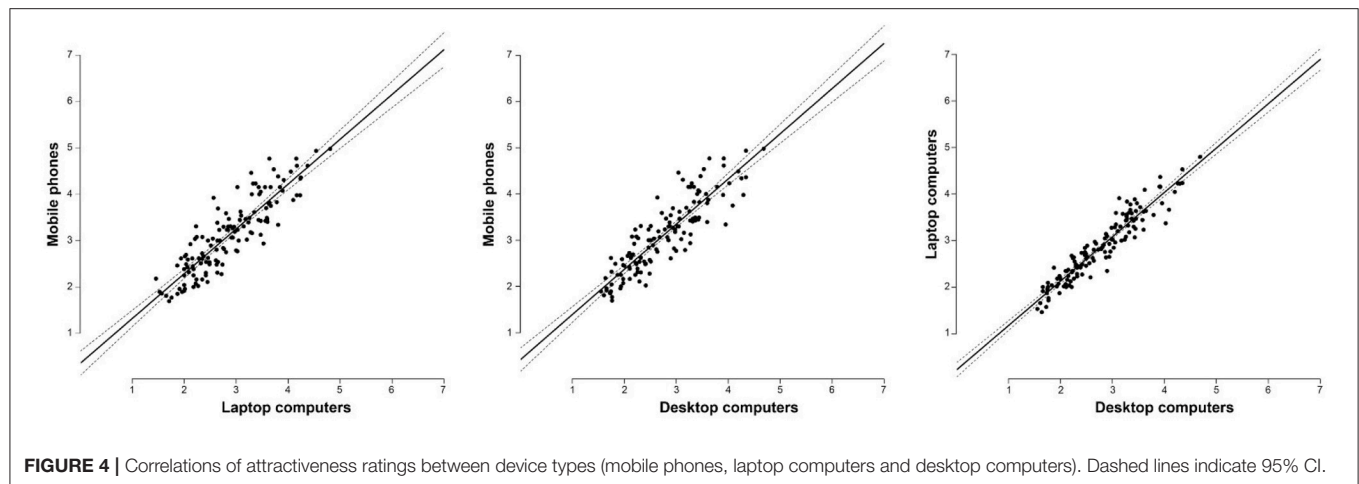


TABLE 3 | Devices and rating descriptive statistics.

		<i>N</i>	Mean attractiveness rating	<i>SD</i>
Device	Mobile phone	116	3.021	0.625
	Laptop computer	232	2.876	0.668
	Desktop computer	136	2.777	0.634

in the mean ratings of formidability and attractiveness. Our results thus showed that ratings based on the three different face views were highly congruent and both perceived formidability and attractiveness ratings appear to be view-invariant. As a subsidiary aim, we have also tested the effect of the device used on attractiveness rating. While there were no differences between ratings performed on desktop and laptop computers, ratings performed using the mobile phones were higher (targets were perceived as more attractive).

Majority of facial perception research uses as stimuli frontal images (Kościnski, 2009), which is in striking contrast with our daily life experience. Moreover, there is a long-standing debate on how human visual system recognizes objects viewed from different angles (Hayward, 2003) and whether object recognition is view-specific, i.e., linked to a specific viewing orientation (Tarr and Bülhoff, 1995), or view-invariant (Biederman and Gerhardstein, 1993). Some evidence suggests that human visual system may be view-specific and process objects differently depending on the viewing angle (Jeffery et al., 2007; but see Jiang et al., 2006). If this were the case, results from perceptual studies



that rely solely on frontal portraits could not be generalized to the other viewpoints. Our results, however, at least when it comes to social perception, do not support this hypothesis.

Our data shows patterns analogous in direction and magnitude to those reported in previous studies that compared assessments based on different stimuli views of both faces and bodies (e.g., frontal \times profile, frontal \times 3D or oblique poses), which likewise showed strong correlations in ratings (Diener et al., 1995; Tovée and Cornelissen, 2001; Valenzano et al., 2006; Davidenko, 2007; Shafiee et al., 2008; Perilloux et al., 2012; Tigue et al., 2012; Dixon et al., 2015; Kościński and Zalewska, 2017). A related study by Tigue et al. (2012) reported that attractiveness of frontal and 3D depictions of women's faces as rated by men were highly correlated ($r = 0.71$) but 3D stimuli received significantly higher mean ratings. Authors suggest that their findings may be an effect of novelty of the 3D visualization. Our study, on the other hand, found no differences between the mean ratings of 2D (frontal and profile) and 3D images. It should be noted, however, that we opted for an alternative to standard 3D visualization. By combining several individual photographs presented in sequence as a spin (360° rotation), we avoided possible bias based simply on differences in capture technology (such as noticeable differences in lighting and colors between 2D and 3D stimuli).

The 360° rotation photographs allowed us to present raters with an all-around view of stimuli heads without running the cost of acquiring 3D capture technologies. Although the resulting visualizations are indeed photorealistic, there is a notable drawback related to implementing this procedure. The capturing and subsequent processing of the images is rather time-consuming and physically demanding, especially for the photographed targets, because one spin takes approx. 5 min and during this time, targets have to sit completely still with fixed gaze, so that controlling for head tilts, yawns, and rolls thus becomes even more critical (Penton-Voak et al., 2001; Hehman et al., 2013; Sulikowski et al., 2015). The use of 3D stimuli captured with actual 3D scanning and 3D reconstructions technology would allow for a variety of target applications including stimuli capture and presentation. For

instance, resulting models could be rotated to arbitrary angles relative to their position during capture, rather than simply displayed in an identical head position in all photographs. Such 3D stimuli would produce more realistic face reconstructions: the main obstacle is the relatively high initial investment into a 3D scanner. Moreover, although 3D facial models are remarkably human-like, they are certainly distinguishable from, and less familiar than, photographs and that could potentially reduce their validity in terms of being a realistic visualization of humans (Crookes et al., 2015). Future studies should investigate whether perception of 3D models differs from 360° rotation photographs.

Interestingly, we found that the device used for viewing the stimuli has a significant influence on the rating. Raters using mobile phones gave on average higher attractiveness ratings than users of laptop or desktop computers. To our best knowledge, no previous study investigated the influence of the device used for viewing on ratings. Although the screen size and resolution of mobile phones are increasing, screen size of handheld devices does limit the size of images that can be viewed on it. That negatively influences the amount of detailed visual information available to the observer, hence potentially limiting the visibility of cues that may affect some aspects of social perception (such as attractiveness). All this may result in ratings higher than those based on viewing images on larger screens which do show more detail. For instance, several studies have reported that more homogenous skin is perceived as more attractive (Jones et al., 2004; Fink et al., 2006; Tsankova and Kappas, 2016; Jaeger et al., 2018; Tan et al., 2018). It is thus possible that lower visibility of such types of imperfections on mobile phones may lead to higher scores. This is a potentially important issue since ever more researchers opt for online data collection. One ought to take into consideration the kind of devices raters decide to use for their viewing and rating, because if a specific subgroup of raters systematically chooses to use a particular kind of device, it could bias the results. In our case, mobile phones were used by nearly one quarter of raters. The results we report are correlations and we just assume that differences in ratings were influenced by

the kind of device raters used. In theory, it is possible that a particular group of raters simultaneously tended to give higher ratings and used mobile phones for viewing and rating. It is likely, however, that the two phenomena are independent of one another, because we found no differences between rater groups in other characteristics.

A potential limitation of our study is the fact that we used a rather specific sample of targets, namely MMA athletes. This fact might limit the generalization of our results. One could expect that MMA fighters would be perceived as highly formidable opponents, as rather specific in appearance (“cauliflower” ears, broken noses, eyebrow scares, etc.), which is why their ratings of formidability and attractiveness might be less variable and/or skewed. In our study, however, raters were not explicitly told that the targets presented to them are MMA fighters. What we found was that mean formidability rating of all three stimuli types on a 7-point scale ≈ 4 (ranging from 2 to 6.2) and skewness of all three stimuli types were between 0.097 and 0.189 (Table 2) and data followed normal distribution. For attractiveness, mean ratings for all three stimuli types were between 2.8 and 2.93 (ranging from 1.7 to 4.8) and skewness between 0.111 and 0.578 (Table 2), hence comparable to average ratings of male facial attractiveness in other studies (e.g., Saribay et al., 2018). It thus seems that the specific nature of our sample does not impede generalization of our finding. Nevertheless, future studies based on less specific samples should further investigate this issue.

To conclude, the findings presented here, along with other recent studies, provide converging evidence that single and multiple view facial images convey highly overlapping information and a single angle view contains enough information about the spatial structural elements of a face to congruently assess formidability and attractiveness, at least in the case of male faces. These results also suggest that studies which use different types of stimulus depiction are, generally speaking, comparable:

this ought to simplify the interpretation of individual studies.

AUTHOR CONTRIBUTIONS

VT, JF, and JH developed the study concept. KK contributed to the study design. Data collection was performed by VT, JF, DS, and ZŠ. VT performed data analysis and interpretation, VT and JF drafted the manuscript, and DS, ZŠ, KK, and JH provided critical revisions. All authors approved the final version of the manuscript for submission.

FUNDING

This research was supported by Czech Science Foundation (GAČR P407/16/03899S), Charles University Research Centre UNCE 204056 and the Sustainability for the National Institute of Mental Health project, grant number LO1611, with financial support from the Ministry of Education, Youth, and Sports of the Czech Republic under the NPU I program.

ACKNOWLEDGMENTS

We thank Klára Coufalová, Radim Pavelka, Tereza Nevolová, Žaneta Slámová, Pavel Šebesta, Dagmar Schwambergová, and other members of Human Ethology group (www.etologiecloveka.cz) for their help with data collection, Mixed Martial Arts Association Czech Republic (MMAA), Anna Pilátová for English proofreading and all volunteers for their participation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02405/full#supplementary-material>

REFERENCES

- Baldassare, M., and Feller, S. (1975). Cultural variations in personal space: theory, methods, and evidence. *Ethos* 3, 481–503. doi: 10.1525/eth.1975.3.4.02a00020
- Berssenbrügge, P., Berlin, N. F., Kebeck, G., Runte, C., Jung, S., Kleinheinz, J., et al. (2014). 2D and 3D analysis methods of facial asymmetry in comparison. *J. Cranio-Maxillofacial Surg.* 42, e327–e334. doi: 10.1016/j.jcms.2014.01.028
- Biederman, I., and Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 1162–1182. doi: 10.1037/0096-1523.19.6.1162
- Brooks, R. C., Shelly, J. P., Jordan, L., and, J. W., Dixon, B. (2015). The multivariate evolution of female body shape in an artificial digital ecosystem. *Evol. Hum. Behav.* 36, 351–358. doi: 10.1016/j.evolhumbehav.2015.02.001
- Bryan, R., Perona, P., and Adolphs, R. (2012). Perspective distortion from interpersonal distance is an implicit visual cue for social judgments of faces. *PLoS ONE* 7:e45301. doi: 10.1371/journal.pone.0045301
- Caharel, S., Jiang, F., Blanz, V., and Rossion, B. (2009). Recognizing an individual face: 3D shape contributes earlier than 2D surface reflectance information. *Neuroimage* 47, 1809–1818. doi: 10.1016/j.neuroimage.2009.05.065
- Calder, A. J., Rhodes, G., Johnson, M., and Haxby, J. (eds.). (2011). *Oxford Handbook of Face Perception*. New York, NY: Oxford University Press.
- Chelnokova, O., and Laeng, B. (2011). Three-dimensional information in face recognition: an eye-tracking study. *J. Vis.* 11:27. doi: 10.1167/11.13.27
- Cooper, E. A., Piazza, E. A., and Banks, M. S. (2012). The perceptual basis of common photographic practice. *J. Vis.* 12:8. doi: 10.1167/12.5.8
- Cornelissen, P. L., Cornelissen, K. K., Groves, V., McCarty, K., and Tovée, M. J. (2018). View-dependent accuracy in body mass judgements of female bodies. *Body Image* 24, 116–123. doi: 10.1016/j.bodyim.2017.12.007
- Crookes, K., Ewing, L., Gildenhuys, J. D., Kloth, N., Hayward, W. G., Oxner, M., et al. (2015). How well do computer-generated faces tap face expertise? *PLoS ONE* 10:e0141353. doi: 10.1371/journal.pone.0141353
- Danel, D. P., Valentova, J. V., Sanchez, O. R., Leongomez, J. D., Varella, M. A. C., and Kleisner, K. (2018). A cross-cultural study of sex-typicality and averageness: correlation between frontal and lateral measures of human faces. *Am. J. Hum. Biol.* 30:e23147. doi: 10.1002/ajhb.23147
- Davidenko, N. (2007). Silhouetted face profiles: a new methodology for face perception research. *J. Vis.* 7:6. doi: 10.1167/7.4.6
- Diener, E., Wolsic, B., and Fujita, F. (1995). Physical attractiveness and subjective well-being. *J. Pers. Soc. Psychol.* 69, 120–129.
- Dixon, B. J., Lee, A. J., Sherlock, J. M., and Talamas, S. N. (2016). Beneath the beard: do facial morphometrics influence the strength of judgments of men's beardedness? *Evol. Hum. Behav.* 38, 164–174. doi: 10.1016/j.evolhumbehav.2016.08.004

- Dixon, B. J., and Rantala, M. J. (2015). The role of facial and body hair distribution in women's judgments of men's sexual attractiveness. *Arch. Sex. Behav.* 45, 877–889. doi: 10.1037/0022-3514.69.1.120
- Dixon, B. J., Vasey, P. L., Sagata, K., Sibanda, N., Linklater, W. L., Dixon, A. F. *Arch. Sex. Behav.* (2011) 40:1271. doi: 10.1007/s10508-010-9680-6
- Dixon, B. J. W., Duncan, M. J. M., and Dixon, A. F. (2015). The role of breast size and areolar pigmentation in perceptions of women's sexual attractiveness, reproductive health, sexual maturity, maternal nurturing abilities, and age. *Arch. Sex. Behav.* 44, 1685–1695. doi: 10.1007/s10508-015-0516-2
- Dunn, T. J., Baguley, T., and Brunsden, V. (2014). From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br. J. Psychol.* 105, 399–412. doi: 10.1111/bjop.12046
- Enlow, D. H., Hans, M. H. G., and McGrew, L. (eds.) (1996). *Essentials of Facial Growth*. Philadelphia, PA: Saunders Saunders.
- Erkelen, C. (2018). Multiple photographs of a perspective scene reveal the principles of picture perception. *Vision* 2:26. doi: 10.3390/vision2030026
- Fink, B., Grammer, K., and Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evol. Hum. Behav.* 27, 433–442. doi: 10.1016/j.evolhumbehav.2006.08.007
- Hall, E. T. (1966). *The Hidden Dimension*. New York, NY: Doubleday and Co.
- Hayward, W. G. (2003). After the viewpoint debate: where next in object recognition? *Trends Cogn. Sci.* 7, 425–427.
- Helman, E., Leitner, J. B., and Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *J. Exp. Soc. Psychol.* 49, 747–754. doi: 10.1016/j.jesp.2013.02.015
- Holzleitner, I. J., and Perrett, D. I. (2016). Perception of strength from 3D faces is linked to facial cues of physique. *Evol. Hum. Behav.* 37, 217–229. doi: 10.1016/j.evolhumbehav.2015.11.004
- Hu, S., Xiong, J., Fu, P., Qiao, L., Tan, J., Jin, L., et al. (2017). Signatures of personality on dense 3D facial images. *Sci. Rep.* 7, 73. doi: 10.1038/s41598-017-00071-5
- Jaeger, B., Wagemans, F. M. A., Evans, A. M., and van Beest, I. (2018). Effects of facial skin smoothness and blemishes on trait impressions. *Perception* 47, 608–625. doi: 10.1177/0301006618767258
- jamovi project (2018). *jamovi (Version 0.9)*. Available online at: <https://www.jamovi.org>
- JASP Team (2018). *JASP (Version 0.9.0.1)*. Available online at: <https://jasp-stats.org/>
- Jeffery, L., Rhodes, G., and Busey, T. (2007). Broadly tuned, view-specific coding of face shape: opposing figural aftereffects can be induced in different views. *Vision Res.* 47, 3070–3077. doi: 10.1016/j.visres.2007.08.018
- Jenkins, R., White, D., Van Montfort, X., and Mike Burton, A. (2011). Variability in photos of the same face. *Cognition* 121, 313–323. doi: 10.1016/j.cognition.2011.08.001
- Jiang, F., Blanz, V., and O'Toole, A. J. (2006). Probing the visual representation of faces with adaptation. *Psychol. Sci.* 17, 493–500. doi: 10.1111/j.1467-9280.2006.01734.x
- Johnston, D. J., Hunt, O., Johnston, C. D., Burden, D. J., Stevenson, M., and Hepper, P. (2005). The influence of lower face vertical proportion on facial attractiveness. *Eur. J. Orthod.* 27, 349–354. doi: 10.1093/ejo/cji023
- Jones, A. L., Kramer, R. S. S., and Ward, R. (2012). Signals of personality and health: the contributions of facial shape, skin texture, and viewing angle. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1353–1361. doi: 10.1037/a0027078
- Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception* 33, 569–576. doi: 10.1068/p3463
- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., and Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum. Behav.* 39, 424–436. doi: 10.1016/j.evolhumbehav.2018.03.008
- Kościnski, K. (2009). Current status and future directions of research on facial attractiveness. *Anthropol. Rev.* 72, 45–65. doi: 10.2478/v10044-008-0015-3
- Kościnski, K., and Zalewska, M. (2017). "Compatibility of facial perception between frontal and profile view," in *4th Annual Conference Polish Society for Human and Evolution Studies*, Krakow.
- Lefevre, C. E., Lewis, G. J., Bates, T. C., Dzhelevova, M., Coetzee, V., Deary, I. J., et al. (2012). No evidence for sexual dimorphism of facial width-to-height ratio in four large adult samples. *Evol. Hum. Behav.* 33, 623–627. doi: 10.1016/j.evolhumbehav.2012.03.002
- Little, A. C. (2014). Facial attractiveness. *WIREs Cogn. Sci.* 5, 621–634. doi: 10.1002/wcs.1316
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. B-Biological Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Maple, J. R., Vig, K. W., Beck, F. M., Larsen, P. E., and Shanker, S. (2005). A comparison of providers' and consumers' perceptions of facial-profile attractiveness. *Am. J. Orthod. Dentofac. Orthop.* 128, 690–696. doi: 10.1016/j.ajodo.2004.09.030
- Meyer-Marcotty, P., Stellzig-Eisenhauer, A., Bareis, U., Hartmann, J., and Kochel, J. (2011). Three-dimensional perception of facial asymmetry. *Eur. J. Orthod.* 33, 647–653. doi: 10.1093/ejo/cjq146
- Miner, M., and Park, D. (2004). A lifespan database of adult facial stimuli. *Behav. Res. Methods Instrum. Comput.* 36, 630–633. doi: 10.3758/BF03206543
- Murphy, K. E. (1994). Preference for profile orientation in portraits. *Empir. Stud. Arts* 12, 1–7. doi: 10.2190/MUD5-7V3E-YBN2-Q2XJ
- Mydlová, M., Dupej, J., Koudelová, J., and Veleminská, J. (2015). Sexual dimorphism of facial appearance in ageing human adults: a cross-sectional study. *Forensic Sci. Int.* 257:519.e1–519.e9. doi: 10.1016/j.forsciint.2015.09.008
- Nomura, M., Motegi, E., Hatch, J. P., Gakunga, P. T., Ng'ang'a, P. M., Rugh, J. D., et al. (2009). Esthetic preferences of european american, hispanic american, japanese, and african judges for soft-tissue profiles. *Am. J. Orthod. Dentofac. Orthop.* 135, S87–S95. doi: 10.1016/j.ajodo.2008.02.019
- Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B. P., Burt, D. M., et al. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proc. R. Soc. B Biol. Sci.* 268, 1617–1623. doi: 10.1098/rspb.2001.1703
- Perilloux, C., Cloud, J. M., and Buss, D. M. (2012). Women's physical attractiveness and short-term mating strategies. *Pers. Individ. Dif.* 54, 490–495. doi: 10.1016/j.paid.2012.10.028
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Rowland, H. M., and Burriss, R. P. (2017). Human colour in mate choice and competition. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160350. doi: 10.1098/rstb.2016.0350
- Rule, N. O., Ambady, N., and Adams, R. B. (2009). Personality in perspective: judgmental consistency across orientations of the face. *Perception* 38, 1688–1699. doi: 10.1068/p6384
- Saribay, S. A., Biten, A. F., Meral, E. O., Aldan, P., Třebický, V., and Kleisner, K. (2018). The Bogazici face database: standardized photographs of Turkish faces with supporting materials. *PLoS ONE* 13:e0192018. doi: 10.1371/journal.pone.0192018
- Sell, A. N., Lukaszewski, A. W., and Townsley, M. (2017). Cues of upper body strength account for most of the variance in men's bodily attractiveness. *Proc. R. Soc. B Biol. Sci.* 284:20171819. doi: 10.1098/rspb.2017.1819
- Shafiee, R., Korn, E. L., Pearson, H., Boyd, R. L., and Baumrind, S. (2008). Evaluation of facial attractiveness from end-of-treatment facial photographs. *Am. J. Orthod. Dentofac. Orthop.* 133, 500–508. doi: 10.1016/j.ajodo.2006.04.048
- Soh, J., Chew, M. T., and Wong, H. B. (2007). An Asian community's perspective on facial profile attractiveness. *Community Dent. Oral Epidemiol.* 35, 18–24. doi: 10.1111/j.1600-0528.2007.00304.x
- Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., et al. (2017). Preferred interpersonal distances: a global comparison. *J. Cross. Cult. Psychol.* 48, 577–592. doi: 10.1177/0022022117698039
- Spyropoulos, M. N., and Halazonetis, D. J. (2001). Significance of the soft tissue profile on facial esthetics. *Am. J. Orthod. Dentofac. Orthop.* 119, 464–471. doi: 10.1067/mod.2001.113656
- Sulikowski, D., Burke, D., Havlíček, J., and Roberts, S. C. (2015). Head tilt and fertility contribute to different aspects of female facial attractiveness. *Ethology* 121, 1002–1009. doi: 10.1111/eth.12412
- Sutherland, C. A. M., Young, A. W., and Rhodes, G. (2017). Facial first impressions from another angle: how social judgements are influenced by changeable and invariant facial properties. *Br. J. Psychol.* 108, 397–415. doi: 10.1111/bjop.12206
- Tan, K. W., Tiddeman, B., and Stephen, I. D. (2018). Skin texture and colour predict perceived health in Asian faces. *Evol. Hum. Behav.* 39, 320–335. doi: 10.1016/j.evolhumbehav.2018.02.003

- Tarr, M. J., and Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1494–1505. doi: 10.1037/0096-1523.21.6.1494
- Thornhill, R., and Gangestad, S. W. (1999). Facial attractiveness. *Trends Cogn. Sci.* 3, 452–460. doi: 10.1016/S1364-6613(99)01403-5
- Thorsten, C. A. (2018). The social psychophysics of human face color: review and recommendations. *Soc. Cogn.* 36, 247–273. doi: 10.1521/soco.2018.36.2.247
- Tigue, C. C., Pisanski, K., O'Connor, J. J., Fraccaro, P. J., and Feinberg, D. R. (2012). Men's judgments of women's facial attractiveness from two- and three-dimensional images are similar. *J. Vis.* 12:3. doi: 10.1167/12.12.3
- Toole, A. J. O., Price, T., Vetter, T., Bartlett, J. C., and Blanz, V. (1999). 3D shape and 2D surface textures of human faces: the role of "averages" in attractiveness and age. *Image Vision Comput.* 18, 9–19. doi: 10.1016/S0262-8856(99)00012-8
- Tovée, M. J., and Cornelissen, P. L. (2001). Female and male perception of female physical attractiveness in front-view and profile. *Br. J. Psychol.* 92, 391–402. doi: 10.1348/000712601162257
- Třebický, V., Fialová, J., Kleisner, K., and Havlíček, J. (2016). Focal length affects depicted shape and perception of facial images. *PLoS ONE* 11:e0149313. doi: 10.1371/journal.pone.0149313
- Tsankova, E., and Kappas, A. (2016). Facial skin smoothness as an indicator of perceived trustworthiness and related traits. *Perception* 45, 400–408. doi: 10.1177/0301006615616748
- Valentine, T., Darling, S., and Donnelly, M. (2004). Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychon. Bull. Rev.* 11, 482–487.
- Valentová, J., Roberts, S. C., and Havlíček, J. (2013). Preferences for facial and vocal masculinity in homosexual men: the role of relationship status, sexual restrictiveness, and self-perceived masculinity. *Perception* 42, 187–197. doi: 10.1068/p6909
- Valentová, J. V., Varella, M. A. C., Havlíček, J., and Kleisner, K. (2017). Positive association between vocal and facial attractiveness in women but not in men: a cross-cultural study. *Behav. Process.* 135, 95–100. doi: 10.1016/j.beproc.2016.12.005
- Valenzano, D. R., Mennucci, A., Tartarelli, G., and Cellerino, A. (2006). Shape analysis of female facial attractiveness. *Vision Res.* 46, 1282–1291. doi: 10.1016/j.visres.2005.10.024

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Třebický, Fialová, Stella, Štěrbová, Kleisner and Havlíček. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Judging Others by Your Own Standards: Attractiveness of Primate Faces as Seen by Human Respondents

Silvie Rádlová, Eva Landová and Daniel Frynta*

RP3 Applied Neurosciences and Brain Imaging, National Institute of Mental Health, Klecany, Czechia

OPEN ACCESS

Edited by:

Ian Stephen,
Macquarie University, Australia

Reviewed by:

Justin Kyle Mogilski,
University of South Carolina
Salkehatchie, United States
Ferenc Kocsor,
University of Pécs, Hungary

*Correspondence:

Daniel Frynta
frynta@centrum.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 August 2018

Accepted: 19 November 2018

Published: 11 December 2018

Citation:

Rádlová S, Landová E and
Frynta D (2018) Judging Others by
Your Own Standards: Attractiveness
of Primate Faces as Seen by Human
Respondents. *Front. Psychol.* 9:2439.
doi: 10.3389/fpsyg.2018.02439

The aspects of facial attractiveness have been widely studied, especially within the context of evolutionary psychology, which proposes that aesthetic judgements of human faces are shaped by biologically based standards of beauty reflecting the mate quality. However, the faces of primates, who are very similar to us yet still considered non-human, remain neglected. In this paper, we aimed to study the facial attractiveness of non-human primates as judged by human respondents. We asked 286 Czech respondents to score photos of 107 primate species according to their perceived “beauty.” Then, we analyzed factors affecting the scores including morphology, colors, and human-likeness. We found that the three main primate groups were each scored using different cues. The proportions of inner facial features and distinctiveness are cues widely reported to affect human facial attractiveness. Interestingly, we found that these factors also affected the attractiveness scores of primate faces, but only within the Catarrhines, i.e., the primate group most similar to humans. Within this group, human-likeness positively affected the attractiveness scores, and facial extremities such as a prolonged nose or exaggerated cheeks were considered the least attractive. On the contrary, the least human-like prosimians were scored as the most attractive group. The results are discussed in the context of the “uncanny valley,” the widely discussed empirical rule.

Keywords: primates, facial attractiveness, visual perception, human preferences, uncanny valley, colors, visual cues

INTRODUCTION

Faces play a key role in the identification of other individuals, which is one of the most important skills needed in social communication of primates (Pascalis and Bachevalier, 1998; Santana et al., 2012). Humans can read emotional expressions from faces and gain a quick insight into the immediate mood of others (Ekman and Friesen, 1986; Fridlund, 1994; Russell, 1994; Calvo and Nummenmaa, 2016). Facial cues also bear information about the individual’s social role (age, sex, and race; for a review, see Yovel, 2016) or personality, such as dominance (Jones et al., 2010), extraversion (Borkenau et al., 2009), trustworthiness (Stirrat and Perrett, 2010), intelligence (Zebrowitz et al., 2002), or emotional stability (Penton-Voak et al., 2006).

Recognition of individual faces is so important that during evolution we gained a complex neural system specialized for just this function (Haxby et al., 2000; Zhao et al., 2018). Because of that, we are able to holistically distinguish faces that subtly differ in minimal position changes of inner facial features, i.e., the eyes, nose, and mouth (Maurer et al., 2002). We also use this recognition ability when evaluating the facial attractiveness: the evaluation is very strict as minimal deviation from the averageness or subtle distinctiveness can be perceived as unattractive or attractive. However, the more different the faces are from our own race and species, the more this ability weakens and diminishes. Using configural processing, humans can process same-race, conspecific faces with a higher success than faces of other races and species (Tanaka et al., 2004; Michel et al., 2006; Ge et al., 2009; Taubert, 2009), and the same applies for non-human primates: according to Gothard et al. (2009), macaques use configural processing when identifying faces of conspecifics, but turn to feature-based mode of analysis when processing pictures of human faces. The way in which humans consider attractiveness of faces of other species thus forms a very interesting question. In this matter, primates represent the perfect group to study—they include species phylogenetically closest to humans with very human-like faces, but also less similar species like the prosimians. Is it possible that human respondents see some primates as human caricatures and evaluate their facial attractiveness using the same facial cues as they use when evaluating facial beauty of conspecifics?

The majority of papers study facial beauty related to sexual attractiveness (for reviews, see Thornhill and Gangestad, 1999; Fink and Penton-Voak, 2002; Kościński, 2007). In this context, the best predictors for facial attractiveness are averageness (Jones and Hill, 1993; Rhodes and Tremewan, 1996; Komori et al., 2009; Zhang et al., 2011), symmetry (Grammer and Thornhill, 1994; Perrett et al., 1999; Scheib et al., 1999), sexual dimorphism (Perrett et al., 1998; Valenzano et al., 2006), smoothness of skin texture and color (Fink et al., 2001, 2006, 2012; Jones et al., 2004; Matts et al., 2007) and an absence of visible defects such as scars (Rankin and Borah, 2003) or congenital face clefts (Tobiasen, 1987).

In studies of human facial attractiveness, many aspects of the preferred facial traits vary under different domain specificity, i.e., different features may be preferred when considering facial attractiveness of short-term sexual partners and long-term romantic partners (DeBruine, 2005), competitors (Fisher, 2004), etc. Although the true nature of domain specificity that lies behind the ranking of primate facial “beauty” is unknown (possibly, the primates may be seen as rivals, cooperators, or may induce care-taking motivation), the respondents hardly evaluate the primates as potential romantic partners. However, recognition of human attractive and unattractive facial features is strongly tied to the identification of healthy and fertile mates and to the increase of one’s fitness (e.g., Thornhill and Gangestad, 1999; Fink and Penton-Voak, 2002; Little et al., 2011). One can thus imagine that the selection pressure led to the perfection of fast and precise ability to assess the attractiveness of the conspecific faces. The question of interest now is whether the cues for the recognition of attractive human faces remain the same when

assessing the attractiveness of non-human, but similar faces. To answer this question, we examined various factors, often described as important cues in the evaluation of human faces, and analyzed their effect on human-perceived beauty of primate faces.

Instead of experimenting with subtle facial details using computer manipulations, we chose to examine the true facial variability of extant primate species, which is enormous. Some primates are more similar to humans than the others, and their facial features—the same features that are heeded in human facial attractiveness, i.e., eyes, nose, mouth, etc.—are often exaggerated to the extremes that may be perceived as human caricatures. And while caricatures (i.e., faces with high distinctiveness, Deffenbacher et al., 1998) may be helpful for a better recognition of individuals (Rhodes et al., 1987; Mauro and Kubovy, 1992), it is the average human faces that are considered as attractive (e.g., Rhodes et al., 2001; Trujillo et al., 2014). Distinctiveness is only seen as attractive when composited from highly attractive features, such as the higher cheek bones, thinner jaws, larger eyes, shorter length between mouth and chin, and between nose and mouth (Perrett et al., 1994).

The sexual dimorphism, i.e., masculinity and femininity, also plays an important role in the perception of human facial attractiveness. For example, Little et al. (2002) found that women showed higher preference for male face masculinity when judging for short-term relationships than when judging for long-term relationships. When studying the effect of sexual dimorphism in non-sexual context, Little et al. (2008) found that both women and men preferred more feminine female faces and more masculine male faces, though the preferences were stronger in women than in men. Other papers (Perrett et al., 1998; Rhodes et al., 2000) report that both women and men preferred more feminine faces, regardless of the face gender. Most of the papers agreed on the lack of difference between male and female respondents in the direction of preferred sexual dimorphism, they only differed in the degree. However, masculine male and female faces are perceived by respondents of both genders as dominant (Perrett et al., 1998). The features that make faces look masculine or feminine are very specific. For example, larger jawbones, more prominent cheekbones, and thinner cheeks are all features of human male faces that differentiate them from female faces (Little and Hancock, 2002). However, the particular features may differ from species to species—e.g., masculine features of a Mandrill rather include elongated jaw and bright colors (Dixon, 2012). Thus, this variable is not fully comparable when studying facial attractiveness across all primates.

Apart from human facial attractiveness, a lot is known about the human-rated attractiveness of animals (e.g., Frynta et al., 2009; Marešová et al., 2009; Landová et al., 2012, 2018; Frynta et al., 2013, 2014; Lišková and Frynta, 2013; Lišková et al., 2015). Specific features, such as an overall shape, body size, achromatic components including pattern, surface (skin/feather/fur) texture and coloration, taxonomic classification and human-likeness, etc., determine whether an animal will be preferred or neglected. As the full variability of primate faces include those that are more and less similar to humans, a mix of factors usually known for affecting the attractiveness of both human faces and animals may

play a role. Thus, many of these factors were included into the analysis.

In short, the purpose of this study is to examine human-perceived attractiveness (i.e., positive affinity toward an object) of primate faces. To our knowledge, this is the first study that focuses on the full variability of primate faces across all taxonomic groups. Other studies so far focused on facial attractiveness of humans or animals that are not closely related to humans (e.g., dogs and cats, Archer and Monton, 2011; Hecht and Horowitz, 2015; foxes, Elia, 2013). With the wide focus of this study, we aim to get insight into the human perception of primates, including the subjective recognition of human-animal boundary.

In our mainly exploratory study, we focused on the following two questions: (1) which factors determine the primate facial attractiveness (or beauty) rated by human respondents? (2) Do these determining factors differ among different primate groups? There are three main groups of primates: the prosimians, which are phylogenetically least related to us, the New World monkeys (Platyrrhini), and its sister taxon Catarrhini, which includes Old World monkeys, gibbons, great apes, and humans. Is the beauty of each of the groups rated using different cues? In search for the answers on these questions, we analyzed the effect of morphology, sexual size dimorphism (SSD), pattern, human-likeness, and colors, and we discussed the findings in terms of known facts about both human facial attractiveness and beauty of animals.

MATERIALS AND METHODS

Selection of Species

There are about 376 known extant primate species (Wilson and Reeder, 2005) covering a wide range of morphological variability. For the purpose of this study, we aimed to choose a number of stimuli that would cover as much variability as possible. Thus, we included at least one species from each genus, except for *Phaner* (Fork-marked Lemur), *Procolobus* (Olive Colobus), *Pseudopotto* (False potto) and *Simias* (Pig-tailed Langur), of which there were no acceptable photographs available at the time of stimuli preparation. We also purposely excluded a human as we did not want to direct the respondents to rank the primate faces in the context of human facial attractiveness. The particular species within genera including similar species were selected based on availability of acceptable photographs. Where there was a high morphological variability within the genus, we included more species (two to eight; e.g., *Cercopithecus*, *Eulemur*, *Macaca*, *Saguinus*, etc.). The East Javan Langur (*Trachypithecus auratus*) was included in both black and orange forms. In case of sexually dimorphic species, only males were included. There is a trade-off between the inclusion of both sexes and taxonomic coverage as the number of stimuli need to be limited so that the respondents stay interested and give reliable rankings. Ten species were represented by two different individuals for a control. The random factors were set in a nested hierarchy. The variance of beauty ranking among individuals of the same species was negligible when compared to that between species (VarCorr function in R: Variance nested in the Group (infraorder/superfamily/family): 0.06898805; Genus: 0.10402177;

Species: 0.24383308; Residual = individual: 0.05087559). We then assessed correlations between the conspecifics. Spearman's correlations for all factors (colors, facial measurements, rankings) were high and significant at the $p < 0.05$ level, except for mouth width and chin (beauty: $r = 0.73$; human-likeness: $r = 0.95$). The dataset contained 117 pictures in total (107 after the removal of the control species/individuals, which were not included in the analyses to avoid pseudo-replication). For the full list of included species, see **Supplementary Appendix 1**.

Preparation of the Stimuli

We collected good quality photographs of primates facing the camera. The main resources were Flickr¹ or Wikimedia Commons² licensed under the Creative Commons license. Supplementary resources were our own photos, photos provided by addressed authors, and books (Rowe, 1996; Mittermeier et al., 2010). For the full list of picture resources, see **Supplementary Appendix 1**.

Each photograph was modified to show the primate face in a standardized position: the background was cut off and set to white and the face (in the form of a bust, see **Figure 1**) was rotated so the eyes were intersecting a straight, notional horizontal line. The primate faces were size-adjusted to cover approximately the same space relative to each other on each image. When there were primates showing an emotional expression (e.g., a smile or a frown) or looking sideways, the photos were retouched so the face showed a neutral expression with eyes looking straight to the camera (see **Figure 2**). Because the used photos could not be standardized under the exact same angle, the primates in the pictures slightly differed in the degree of rotation on both vertical and horizontal axes. Thus, we could not test the effect of symmetry on human rankings of primate facial "beauty" as it clearly corresponded to the rotation of the faces. This rotation had no effect on any of the explained variables (none of the Spearman's correlations were significant at the $p < 0.05$ level) and thus was excluded from further analyses.

Definition of the Groups

The recognized taxonomy of Primates consists of seven distinct groups: Lorisoidea (African and Asian prosimians), Lemuroidea (Madagascar prosimians), the Tarsiens, Platyrrhini (New World primates), Cercopithecoidea (Old World monkeys), hylobatids (gibbons), great apes, and humans (Purvis, 1995; Yoder and Irwin, 1999; Pastorini et al., 2001; Geissmann et al., 2004; Mayor et al., 2004; Lei, 2008; Finstermeier et al., 2013). To identify groups of reasonable morphological variability suitable for the purpose of the analysis of human-rated facial attractiveness of primates, we performed the canonical variate analysis (CVA) using the geometrical morphometry data (see below, Section "Shape"). The CVA separated the primates into three distinct groups (see **Figure 3**), referred to as prosimians, Platyrrhini and Catarrhini in the manuscript. The analysis also confirmed morphological distinctiveness of humans when compared to other primates.

¹<https://www.flickr.com/>

²<https://commons.wikimedia.org>



FIGURE 1 | Examples of the stimuli rated by the respondents. The depicted primates are, from upper left to upper right: Pygmy Marmoset (*Cebuella pygmaea*), Philippine Tarsier (*Tarsius syrichta*), and Goeldi's Monkey (*Callimico goeldii*); from lower left to lower right: Red Slender Loris (*Loris tardigradus*), Black Lemur (*Eulemur macaco*), and Northern Talapoin Monkey (*Miopithecus ogouensis*).

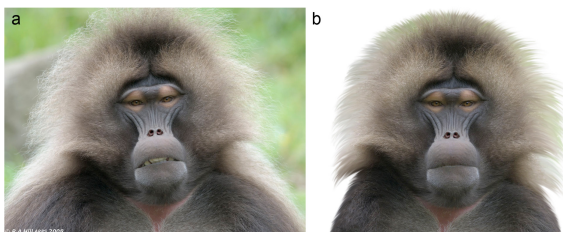


FIGURE 2 | Example of a standardization/modification of the stimuli. **(a)** The original, unaltered picture of a Gelada (*Theropithecus gelada*). **(b)** We modified the photo to use it as a stimulus: the background was cut out, the head was rotated to a straight vertical position and the mouth was closed. Photo© Alan Hill, used with a permission.

Testing Human Preferences

Preferences for each of the primate faces were assessed using an online survey following Frynta et al. (2010) and Lišková and Frynta (2013). The respondents ($n = 286$, 91 men and 199 women) were Czech citizens, 15–69 years old (mean age was 22.7 years). Their task was to rate each of the faces on a scale (1–7 Likert scale; 1 = the most “beautiful,” 7 = the least “beautiful” or “ugly”) according to their perceived “beauty.” The photographs, resized to 360×540 pixels, were presented one by one on a computer screen in a random order. Prior to the presentation of the stimuli, the respondents were able to see the whole variability of the stimuli in the form of thumbnail-sized preview pictures (160×240 pixels). After that, the respondents started to score the pictures. The whole set was divided into groups of 39 photos, and after evaluating each of the groups, the respondents were

allowed to take a rest, although the majority of the respondents finished the scoring without the need of a break. In total, all 286 respondents rated all 117 pictures.

Explanatory Variables

Shape

In our study, we aimed to cover the whole facial variability including the length of facial hair and the forehead size of animal (primate) faces. However, landmarks usually used in human facial studies either only include the shape and position of the eyes, nose, mouth, and chin (e.g., Mitteroecker et al., 2013), or include landmarks that are not applicable for frontal view of primate faces (e.g., Sforza et al., 2007). Thus, we adopted the landmarks of Borgi et al. (2014), who already defined landmarks of animal faces, which, with a few modifications, could easily fit our experimental stimuli (see **Figure 4**): (A) top of the head, (B) right side of the face, (C) left side of the face, (D) end of chin, (E1, G1) outer sides of right and left eyes, respectively, (E2, G2) inner sides of right and left eyes, respectively, (F) middle point of the reference cross, (H) right side of the nose, (I) left side of the nose, (J) tip of the nose, (K) left end of the mouth, (L) middle point of the mouth crossing the reference line, (M) right end of the mouth, (N) top point of head hair, (O1, O2) right and left tips of side hair, (P) tip of the chin hair (beard). Five human facial measurements were added (photos were selected randomly from the FEI Face Database; Thomaz, 2012) and these data were then used to perform the CVA analysis (see above in Section “Definition of the Groups”).

The landmarks were then converted to traditional morphometric variables: the face height (the A–D distance

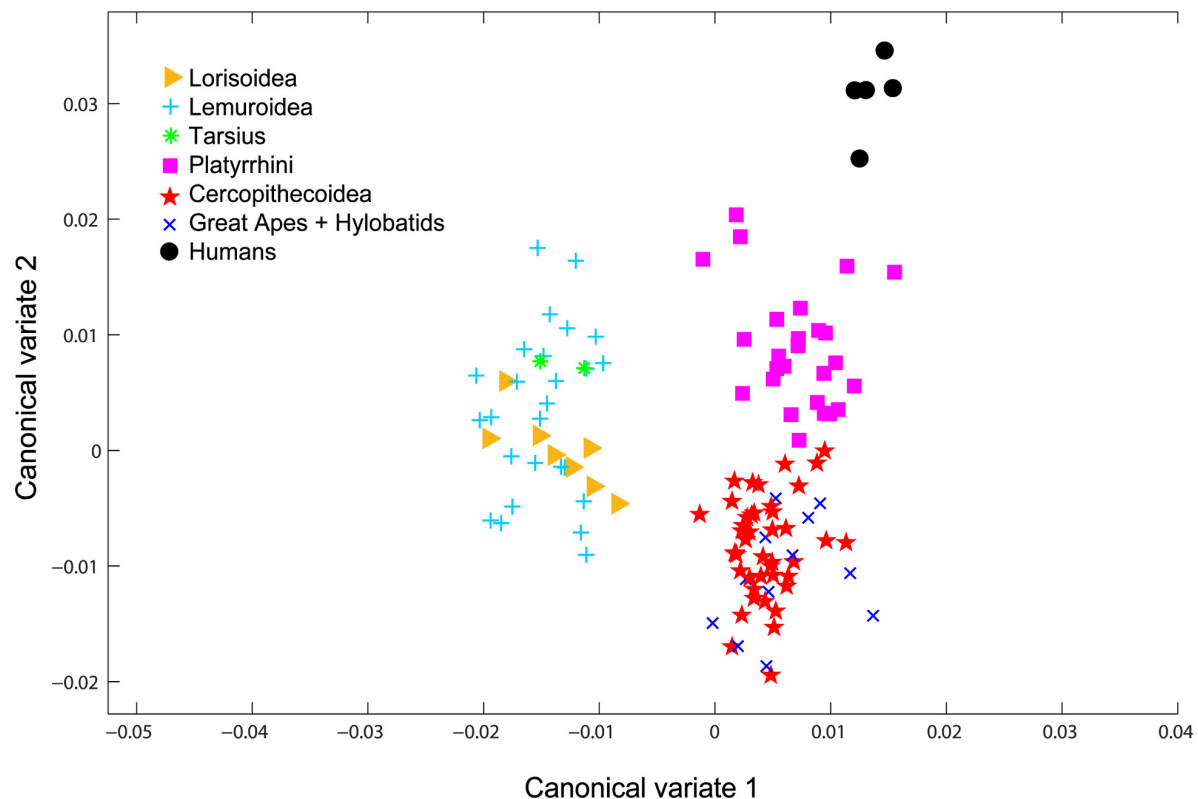


FIGURE 3 | CVA analysis of the seven phylogenetic groups of primates. The analysis is based on the geometrical morphometry data measured from the primate photos and shows grouping into three main primate groups: the prosimians, Platyrrhini, and Catarrhini. Humans also form a separate group, but were excluded from the analyses.

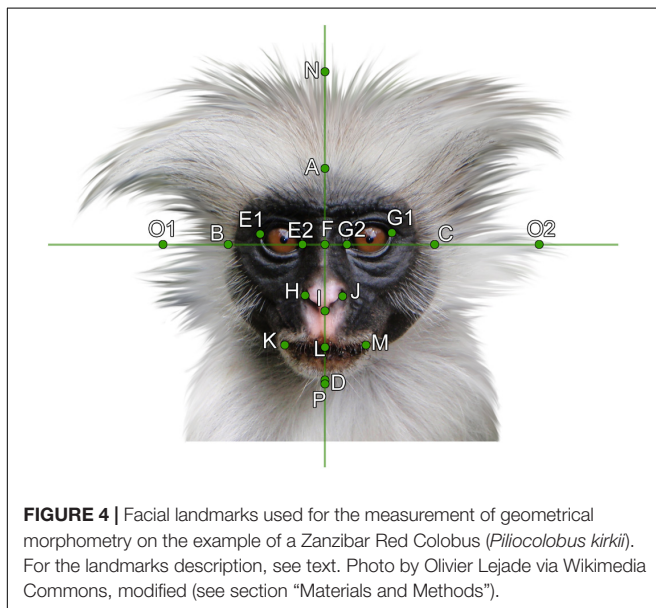


FIGURE 4 | Facial landmarks used for the measurement of geometrical morphometry on the example of a Zanzibar Red Colobus (*Ptilocolobus kirkii*). For the landmarks description, see text. Photo by Olivier Lejade via Wikimedia Commons, modified (see section “Materials and Methods”).

measured in pixels) and width (B–C), forehead height (A–F), eyes size (averaged E1–E2 and G1–G2 distances), nose length (F–I) and width (H–J), mouth width (K–M), side-hair (averaged

O1–B and C–O2), top-hair (N–A), beard (D–P), interocular length (E2–G2), eyes-to-mouth distance (F–L), philtrum (nose-to-mouth distance, I–L), and chin (L–D). All analyzes and data transformations involving the landmarks were done using the IMP software series (Zelditch et al., 2012).

We then extracted maximum likelihood factors from these traits (varimax normalized) to reduce the number of morphological factors for the GLM/GLS analyses and especially to eliminate mutually correlated variables. The first extracted factor, accounting for 20.5% of variation, was interpreted as “outer facial features” (the height of the face and forehead on one side and the length of the beard and top-hair and width of the side-hair on the other side of the axis), while the second one (22.5%) corresponded to “inner facial features” (mainly the distance between eyes and nose from the mouth on one side and the size of the eyes and their distance on the other side of the axis; for factor loadings, see Figure 5).

Colors

To examine the effect of colors on the respondents’ ranking, we used the software Barvocuc (Rádlová et al., 2016) to extract specific information about hues, lightness and saturation of each of the stimulus pictures converted to the HSL colorspace. For a detailed description of the Barvocuc software, see (Lišková and Frynta, 2013 and Lišková et al., 2015). The variation in color is

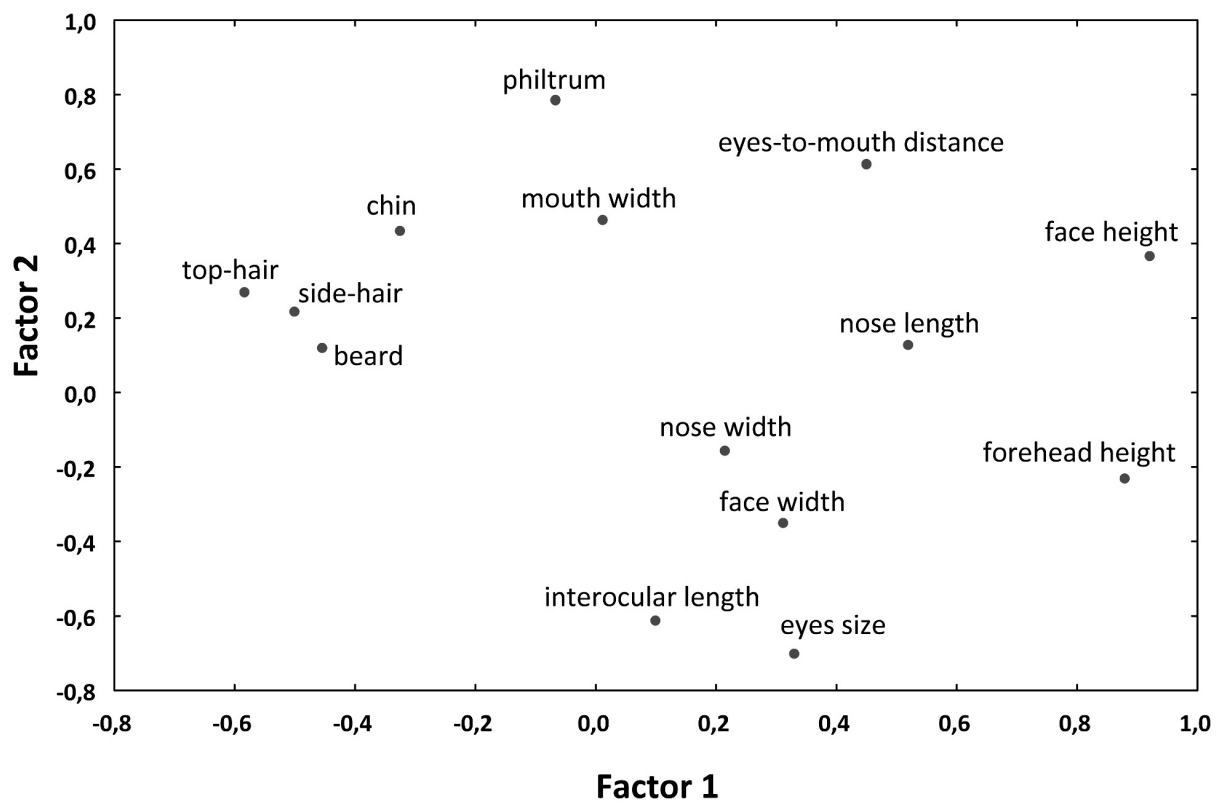


FIGURE 5 | Plot of loadings of varimax normalized maximum likelihood factors, computed from morphometric traits of the primate pictures. The first extracted factor accounts for 20.5% of variation and can be interpreted as “outer facial features.” The second factor accounts for 22.5% of variation and corresponds to the “inner facial features.”

much smaller among primates than other animals, such as birds. Thus, the picture set included only the following colors, which we pre-defined using the software to describe the primate faces as accurately as possible: red (corresponding to the reddish brown in the primate photos) $<350^\circ$; 18°), orange $<18^\circ$; 45°), yellow (corresponding to yellowish brown) $<45^\circ$; 75°), and bluish tint $<170^\circ$; 270°). The variability of blue color was too low in the dataset (only two primates possessed true blue facial parts: the Mandrill *Mandrillus sphinx* and Golden Snub-nosed Monkey *Rhinopithecus roxellana*). However, the blue color was present in a small amount on several photographs in the form of a bluish tint. Because blue color plays a crucial role in the determination of human preferences toward many groups of animals (Frynta et al., 2010; Lišková and Frynta, 2013; Lišková et al., 2015; Ptáčková et al., 2017), we decided to include the “bluish tint” color (blue hue minus the facial parts of *M. sphinx* and *R. roxellana*) as an explanatory variable for the analysis.

The values for saturation (S) and lightness (L) covered the interval 0–1. We defined three additional colors: black ($L < 0.20$), white ($L > 0.71$), and gray ($S < 0.15$). The white background of the stimuli was set to transparent and thus excluded from the calculation. In order to improve normality, the portion of colored pixels in the tested pictures was square-root arcsin transformed prior to the analyses. We also included the “pattern,” computed using an edge detection method (Sobel, 1978) as an explanatory

variable in the analyses. The highest values of the pattern variable corresponded to the agouti coloration of some of the primates.

Sexual Size Dimorphism

Sexual dimorphism as studied in context of human facial attractiveness usually refers to the sexual shape dimorphism. This variable, however, is hardly comparable to primate facial sexual shape dimorphism. It is because each face represents a different species, and the particular features shaping males and females may differ for each species. These features may include various characteristics such as conspicuous cheeks, enlarged noses, colorful prolonged snouts, etc. These masculine features are not directly comparable to those of human males, which are defined by, e.g., subtle changes in jawbones size or more prominent cheekbones, as mentioned above (Little and Hancock, 2002). However, it is possible to use related species characteristics that is available from published sources—the sexual dimorphism in body size. Sexual selection, alongside with the increase in male body size, promotes the emergence of novel conspicuous traits, including those visible on primate faces. Thus, the larger the size difference between the sexes, the larger are the distinctive facial features. For example, size dimorphism in canines increases with SSD in primates (Leutenegger and Kelly, 1977; Kay et al., 1988) and thus modifies the primate mouth shape as bigger canines require more elongated jaws (Weston et al., 2004). Adult

males of sexually dimorphic (male-larger) species display red and blue sexual skin (e.g., the Mandrill), capes of hair, and facial adornments (e.g., the Bald Uacari, Proboscis Monkey, Golden Snub-nosed Monkey, or the Orangutans; Dixon, 2012). In this paper, we utilized this variable to indirectly examine the effect of conspicuous traits on the human evaluation of primate facial attractiveness.

Sexual size dimorphism was expressed using the Lovich and Gibbons ratio (LG ratio; Lovich and Gibbons, 1992), which produces measures of sexual dimorphism continuous around 0. The values were computed as follows: (body weight of the larger sex/body weight of the smaller sex) -1, negative by convention when males are the larger sex and positive when females are larger than males. LG ratios of the primates set varied within the range of -1.371 in the Western Gorilla (*Gorilla gorilla*) to 0.313 in Lemurine Night Monkey (*Aotus lemurinus*). The body weights were adopted from Lindenfors and Tullberg (1998); Gordon (2006), and Mitteroecker et al. (2013).

Human-Likeness

Sixty respondents (different from the ones evaluating the attractiveness) repeated the procedure described above (Section “Testing Human Preferences”) to rate the primates’ human-likeness (1–7 Likert scale; 1 = most human-like, 7 = not human-like at all). Agreement among the respondents in human-likeness of the primates was exceptionally high. The intra-class correlation (ICC, see later in the text), assessed using a two-way, consistency measure, was in an excellent range: ICC = 0.986 for average-measure, 0.553 for single-measure (Hallgren, 2012). To ensure that the knowledge of the great apes being the most phylogenetically related to humans did not distort the overall agreement, we also checked for the ICC of the data excluding the Homioidea (great apes and gibbons): ICC was 0.983 for average-measure, 0.5 for single-measure; i.e., these analyses show that the respondents agreed well on the human-likeness of the particular primate groups/species and their rankings were not influenced by just the most human-like apes. The multivariate analysis of variance revealed no effect of gender, age, nor their interaction. Thus, we pooled the dataset and used the mean values in the analyses as a reliable estimate of human-likeness of the ranked primate species.

Statistical Analyses

In order to quantify and test congruence in species ranking provided by different respondents, we adopted a two-way, consistency, average-measures intra-class correlation (ICC; McGraw and Wong, 1996; Hallgren, 2012) computed in R (irr package). Principal component analysis (PCA) was performed to visualize the multivariate structure of the data sets and to extract uncorrelated axes for further analyses. MANOVA and General Linear Models (LMs) were applied to test the effects of independent explanatory variables. Full LMs were further reduced according to Akaike criterion until log-likelihood tests revealed a significant comparison between the full and reduced models. Mann–Whitney test was used as a non-parametric alternative for variables deviating from normality (raw scores). The contribution of the explanatory variables (constraints) to

the attractiveness rating of the primate faces was examined in redundancy analysis (RDA) as implemented in the R package vegan (Oksanen et al., 2017). RDA is a multivariate direct gradient method. It extracts and summarizes the variation in a set of response variables (subjective evaluation of primate beauty) that can be explained by a set of explanatory variables (see Section “Explanatory Variables”). This analysis permits to plot both response and explanatory variables to a space defined by the extracted gradients and enables to detect redundancy (i.e., shared variability) between sets of response and explanatory variables. Statistical significance of the gradients was confirmed by permutation tests. Most of the calculations were performed in R (R Development Core Team, 2010) and Statistica 9.1. (StatSoft, 2010).

Ethics Statement

This study was carried out in accordance with the recommendations of Institutional Review Board (IRB), Faculty of Sciences, Charles University in Prague, approval no. 2013/7, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the IRB.

RESULTS

Agreement Among the Respondents

Results of the ranking procedure revealed considerable congruence among the respondents. Although the reliability of the individual rankings was quite low (ICC = 0.147, 0.204, 0.182 for men, women, and pooled data, respectively, with all $p < 0.001$), the ICC for the average-measures was in an excellent range: ICC = 0.940 for men, 0.981 for women, and 0.985 for the pooled data (Shrout and Fleiss, 1979; Cicchetti, 1994). These results indicate that there was a high degree of agreement within the group of the respondents and suggest that preferences for primate faces were rated similarly. Also the correlation between ranks provided by male and female respondents was very high: $r^2 = 81.6\%$, $p < 0.001$. Multivariate analysis of variance revealed no effect of age (Wilks = 0.6191, $F_{171, 109} = 0.97$, $p = 0.575$) or age \times gender interaction (Wilks = 0.5823, $F_{171, 109} = 1.13$, $p = 0.2436$), nevertheless, a small, but significant effect of gender (Wilks = 0.5384, $F_{171, 109} = 1.35$, $p = 0.041$) was found. To identify the species that substantially contributed to the gender differences, we performed Mann–Whitney U tests comparing the raw ranks of each species in male and female respondents; the levels of significance were Bonferroni-corrected. Men significantly differed in their preferences from women in only five cases, all of which were preferred by men more than by women: de Brazza’s Monkey (*Cercopithecus neglectus*), Patas Monkey (*Erythrocebus patas*), Humboldt’s Woolly Monkey (*Lagothrix lagotricha*), Drill (*Mandrillus leucophaeus*), and the Sumatran orangutan (*Pongo abelii*). Because the gender differences were small and involved only 5 out of 107 examined species of primates, we decided to pool the genders in further analyses concerning the means or multivariate axes (RDA) computed from the preference ranks. Both of these methods

extract the agreement among respondents and thus further blend the minor effect of gender.

The Attractiveness

The primates whose faces were rated as the most “beautiful” were mostly prosimians: the top winners were the Black-and-white Ruffed Lemur (*Varecia variegata*), Ring-tailed Lemur (*Lemur catta*) and Southern Lesser Galago (*Galago moholi*). Moreover, little monkeys such as the marmosets (Callitrichinae) were favorite among the respondents, together with apes such as the Agile Gibbon (*Hylobates agilis*) and Bonobo (*Pan paniscus*). In contrast, the Proboscis Monkey (*Nasalis larvatus*) or the Bald Uacari (*Cacajao calvus*) were rated as the “least beautiful” (or “ugly”).

When overviewed within the particular groups, all three groups included both attractive and unattractive species (see **Figure 6**). However, the prosimians were rated as significantly more attractive than the other groups (*post-hoc* Tukey test, $p < 0.01$). The particular cues affecting the respondents’ decision and the relation to the uncanny valley theory is discussed below in the respective sections.

RDA Analysis of the Factors Affecting Attractiveness

We employed RDA to examine the contribution of various explanatory variables to the ratings of primate facial attractiveness. We utilized the automatic model-building feature based on both Akaike criterion (but with permutation tests) and on permutation p -values. Both methods agreed on the inclusion of the following variables into the reduced model, which were then confirmed as significant by the sequential “Type I” test (n permutations = 10,000): Factor2 (i.e., inner facial parts; $F_{1,100} = 22.4244$, $p < 0.0001$), human-likeness ($F_{1,100} = 6.0119$, $p < 0.0001$), blue color ($F_{1,100} = 3.4310$, $p = 0.0011$), Factor1 (i.e., outer facial parts; $F_{1,100} = 2.8690$, $p = 0.0050$), LG ($F_{1,100} = 2.3811$, $p = 0.0125$), and pattern ($F_{1,100} = 1.9238$, $p = 0.0400$). The RDA model has generated six constrained axes, which explained 28.08% of the full variability.

The visualization of the RDA results (see **Figure 7**; note that for a better clarity, we multiplied human-likeness and LG by -1 so the higher the number, the higher is both the human-likeness and exaggeration of the male facial parts) showed that Factor2, i.e., the inner facial parts, dominated the first multivariate axis (RDA1; correlation of RDA1 site scores with Factor2: $r^2 = 72.2\%$, $p < 0.0001$). As the most attractive species are located on top and the least attractive on the bottom of the graph (second axis), we can conclude that the RDA2 axis corresponds to the actual attractiveness of the species. Correlation of the mean attractiveness scores with the RDA2 site scores supports this: $r^2 = 69.6\%$, $p < 0.0001$. The only factors associated with this attractiveness irrespective of the second axis (and thus the primate grouping, which corresponds to this axis) are blue color (positive effect) and pattern (negative effect). The graph clearly shows that the grouping of the primates (based on real morphology) reflects the respondents’ ratings of the species’ beauty, i.e., the respondents’ classification of the primate facial beauty differs among the groups and is mainly based on the inner facial properties of the species. Both the extent of human-likeness

and the extent of male sexual dimorphism (-LG) feed this second morphological axis.

GLM of the Factors Affecting Attractiveness

In order to examine which factors contribute to the variability of preference rankings, we performed LMs (see **Table 1**). The initial full model of all the primates together ($n = 107$) included the group, outer (Factor1) and inner (Factor2) facial features, LG, human-likeness, mean lightness, pattern, mean saturation, reddish brown, orange, yellowish brown, and bluish tint. After reduction using the Akaike Information Criterion (AIC, Akaike, 1998), the reduced model explained 34.4% of variation in preference ranks ($p < 0.0001$) and included the group $F_{2,100} = 11.0290$, $p < 0.001$, inner facial features ($F_{1,100} = 16.2674$, $p = 0.0001$), LG ($F_{1,100} = 5.7186$, $p = 0.0187$), human-likeness ($F_{1,100} = 8.0299$, $p = 0.0056$), and bluish tint ($F_{1,94} = 9.4351$, $p = 0.0027$).

We then conducted the same analysis separately for each of the groups Catarrhini, Platyrrhini, and prosimians (see **Table 1b,c**). Catarrhini ($n = 50$): the reduced model ($r^2 = 54.6\%$, $p < 0.0001$) included the inner facial features ($F_{1,44} = 27.1710$, $p < 0.0001$), LG ($F_{1,44} = 5.5832$, $p = 0.0226$), human-likeness ($F_{1,44} = 16.6525$, $p = 0.0002$), reddish brown, and bluish tint ($F_{1,44} = 12.4338$, $p = 0.0010$). Platyrrhini ($n = 24$): the LG, human-likeness, mean lightness, pattern, orange, and yellowish brown color remained in the model ($r^2 = 40.1\%$, $p < 0.0001$), but only the mean lightness ($F_{1,17} = 6.8828$, $p = 0.0178$) and yellowish brown ($F_{1,17} = 6.4430$, $p = 0.0212$) retained significance. The model for prosimians ($n = 33$) failed to explain any variability and was not significant.

DISCUSSION

The Effect of Shape

The inner facial parts represent one of the strongest factors determining the beauty of the species within the group of the primates most similar to us, i.e., the Catarrhini. Thus, the size of eyes, interocular length, mouth width, and length from the nose to the mouth (or eyes to the mouth) are strong cues that our respondents use as a guide when ranking the “beauty” of Catarrhine faces. Consequently, although the respondents were not instructed to categorize the primates (the scoring procedure in our experiment instructed the respondents to assign scores to the pictured faces according to the subjectively perceived beauty), the RDA2 axis shows that the respondents still categorized the ranked subjects, and this categorization was mainly based on the inner facial parts of the primate faces. This phenomenon is often reported in studies focused on human perception of animal beauty (e.g., snakes; Marešová et al., 2009; Landová et al., 2012; birds; Lišková et al., 2015) and resembles the task recognized as unsupervised human categorization (Pothos and Chater, 2002; Pothos and Close, 2008).

In literature, the understanding of the role of inner and outer facial features is unclear. Some authors claim that young children mostly use the outer facial features as the cues for facial recognition, and then this pattern switches to the “adult version,” in which the faces are recognized using the inner

TABLE 1 | The final reduced models (GLM) describing the effect of the morphology, colors, and human-likeness on the attractiveness scoring of each of the main primate groups.

	ANOVA			Coefficients			
	<i>Df</i>	<i>F</i>	<i>Pr (>F)</i>	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>Pr (> t)</i>
(a) All primates (attractiveness, $r^2 = 0.3437$)							
(Intercept)				2.7457	0.4545	6.0420	0.0000
Group	2	11.0290	0.0000				
Group (Platyrrhini)				0.2927	0.1742	1.6800	0.0961
Group (prosimians)				−0.2936	0.2743	−1.0710	0.2869
Factor2	1	16.2674	0.0001	0.3488	0.1238	2.8180	0.0058
LG	1	5.7186	0.0187	−0.5119	0.2198	−2.3290	0.0219
Human-likeness	1	8.0299	0.0056	0.2141	0.0987	2.1690	0.0325
Bluish tint	1	9.4351	0.0027	−1.5962	0.5196	−3.0720	0.0027
Residuals	100						
(b) Catarrhini (attractiveness, $r^2 = 0.5461$)							
(Intercept)				2.5497	0.3633	7.0170	0.0000
Factor2	1	27.1710	0.0000	0.4386	0.1109	3.9540	0.0003
LG	1	5.5832	0.0226	−0.2683	0.1959	−1.3700	0.1777
Human-likeness	1	16.6525	0.0002	0.2576	0.0819	3.1450	0.0030
Reddish brown	1	2.1168	0.1528	0.8435	0.5552	1.5190	0.1359
Bluish tint	1	12.4338	0.0010	−1.4408	0.4086	−3.5260	0.0010
Residuals	44						
(c) Platyrrhini (attractiveness, $r^2 = 0.4006$)							
(Intercept)				5.2577	1.5350	3.4250	0.0032
LG	1	3.4340	0.0813	−1.4253	0.7377	−1.9320	0.0702
Human-likeness	1	0.2547	0.6203	−0.7377	0.3366	−2.1920	0.0426
Mean lightness	1	6.8828	0.0178	3.7882	1.2575	3.0130	0.0078
Pattern	1	3.6745	0.0722	2.9365	1.5838	1.8540	0.0812
Orange	1	0.6811	0.4206	0.5885	0.4672	1.2590	0.2249
Yellowish brown	1	6.4430	0.0212	−1.7697	0.6972	−2.5380	0.0212
Residuals	17						

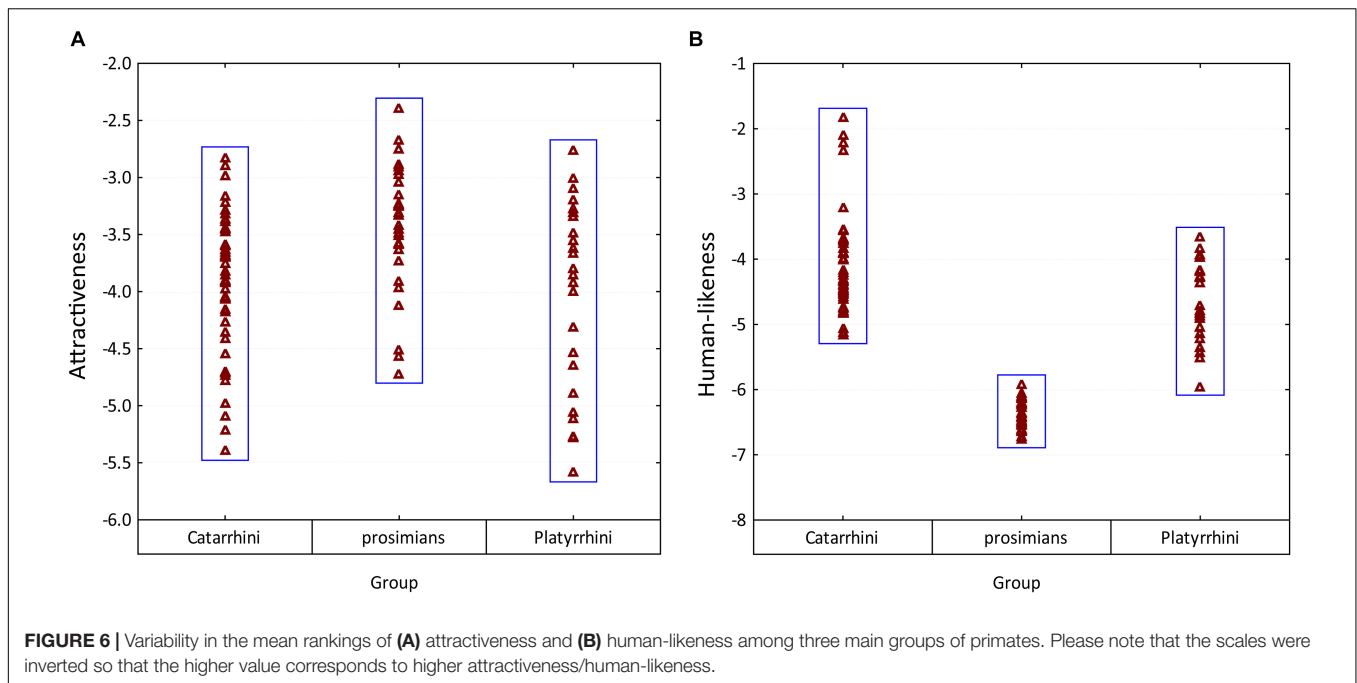
features (Campbell et al., 1995; Turati et al., 2006). Other authors argue that both children and adults use inner facial features for the recognition of familiar faces, but outer facial features for recognition of the unfamiliar ones (Ellis et al., 1979; Want et al., 2003; Bonner and Burton, 2004; Ge et al., 2008). Our results show that the inner facial features are not only used for categorization of the primates, but also play a very important role in the assessment of the facial beauty of the Catarrhine primates. Outer facial features are used to a much less extent, but also appear to contribute to the assessment of primate beauty (see Figure 7).

Colors and Pattern in Primate Facial Attractiveness

Our results show that two colors affect the attractiveness of primate faces: the bluish tint (in Catarrhines and the full picture set) and the yellowish brown color (Platyrrhines). In literature, colors do play a role in the assessment of attractiveness, especially the red color, which is important for both humans and non-human primates. Human faces exhibiting brighter red are perceived as healthier and more attractive (Re et al., 2011). Female Rhesus Macaques prefer males with redder faces (Waite et al.,

2003; but see Waite et al., 2006, where this preference only applied to red hindquarters). Moreover, red clothing or even extraneous red (for example, red background of a presented picture stimulus) is perceived as more attractive by both human respondents (Elliot et al., 2010) and non-human primates (Hughes et al., 2015).

We examined the full variability of colors present in the picture set of primates, i.e., not only red, but also orange, yellow, and the bluish tint. Within our examined picture set, only three primates possessed bright red coloration of the face (The Bald Uacari, Silvery Marmoset, and Japanese Macaque). Thus, we instead tested the effect of the overall presence of the red color, mostly expressed as darker red or reddish brown fur color. However, we found no effect of this color on human preferences. The only color that positively affected human decisions toward all primates (regardless of the particular groups) was blue—the same color that is, within the context of facial attractiveness, usually perceived negatively, as blue, pale faces indicate low oxygenation and poor health (Stadie, 1919). However, blue is very often reported as the most preferred color in other studies examining human rating of animal beauty, e.g., parrots (Frynta et al., 2010), birds (Lišková and Frynta, 2013; Lišková et al., 2015), and even snakes (Ptáčková et al., 2017). In the latter case, the color was present only in the form of a bluish tint. Similarly, in this



paper, this bluish tint affected the overall evaluation of the beauty and was rather the effect of the photos than the primate species themselves.

When analyzing the particular groups, the GLM revealed that the attractiveness of the New World monkeys was positively affected by the yellowish brown color. This is in agreement with the perception of human faces as more attractive (Stephen et al., 2012) and healthier (and thus more attractive: Stephen et al., 2009) show that the respondents increased yellow color together with red and overall lightness when aiming to create a healthy-looking human face. Even papers dealing with animal attractiveness report preference for yellow color in some cases, especially when rating the beauty of birds (Lišková and Frynta, 2013; Lišková et al., 2015). However, animal attractiveness is usually mainly determined by the pattern and achromatic contrasts (Lišková and Frynta, 2013; Lišková et al., 2015). Similarly, the only other variable next to the yellow color that explained the attractiveness ratings of the New World primates was the mean lightness—the respondents rated darker monkeys as more beautiful. This agrees with the animal studies but contrasts human-facial studies in which either lighter (Van den Berghe and Frost, 1986; Stephen et al., 2009; Coetzee et al., 2012) or medium-toned, but not pale or black faces are rated as more attractive (Frisby, 2006; Stephen et al., 2012, but see Fink et al., 2001).

Although the pattern variable was dropped out from the final LMs, the multivariate analysis shows that, at least to some extent, it negatively affects the evaluation of overall primate facial attractiveness. This seems to contradict some of the papers that report positive effect of pattern to the evaluation of animal (Lišková et al., 2015; Ptáčková et al., 2017) and mammalian (Landová et al., in prep.) beauty. However, it is the highly contrasting pattern of large spots, stripes, and other marks, that

positively affects the perceived animal beauty; not the diminutive unevenness of the fur color (i.e., agouti-type fur coloration), which is what the pattern variable used in this study corresponded to and which was perceived negatively. The primates possessing contrasting patches of fur coloration (e.g., the Black and White Ruffed Lemur and the Ring-tailed Lemur) were still considered the most beautiful.

Homogenous skin color distribution and surface topography (wrinkles), signs of health and age, also affect human preferences for attractive conspecific faces (Samson et al., 2010) and could possibly affect the preferences of primate faces as well. However, our set of stimuli was not controlled for the age of the depicted individuals (they were all adults of unspecified age) and thus we could not test the effect of the features affecting perception of age. Moreover, majority of the species included in the study possessed a face that was fully covered by fur. A carefully designed experiment with more uniform stimuli varying only in facial surface topography (e.g., faces of chimpanzees of varying age, or a manipulated picture set) would be needed to examine this interesting question.

SSD, Averageness, and Facial Extremities

There is not much variability within the prosimians in SSD, as most species have genders of similar size (Dixon, 2012). Platyrrhine primates differ more (variance in LG reaches from -0.72 in the Brown Capuchin to 0.31 in the Lemurine Night Monkey), but still lack the most prominent facial extremities typical for many male-larger Catarrhine species such as the Western Gorilla, Mandrill, Drill, Orangutans, Golden Snub-nosed Monkey, Proboscis Monkey, Patas Monkey, Gelada, or the Lion-tailed Macaque. Thus, it is not surprising that the

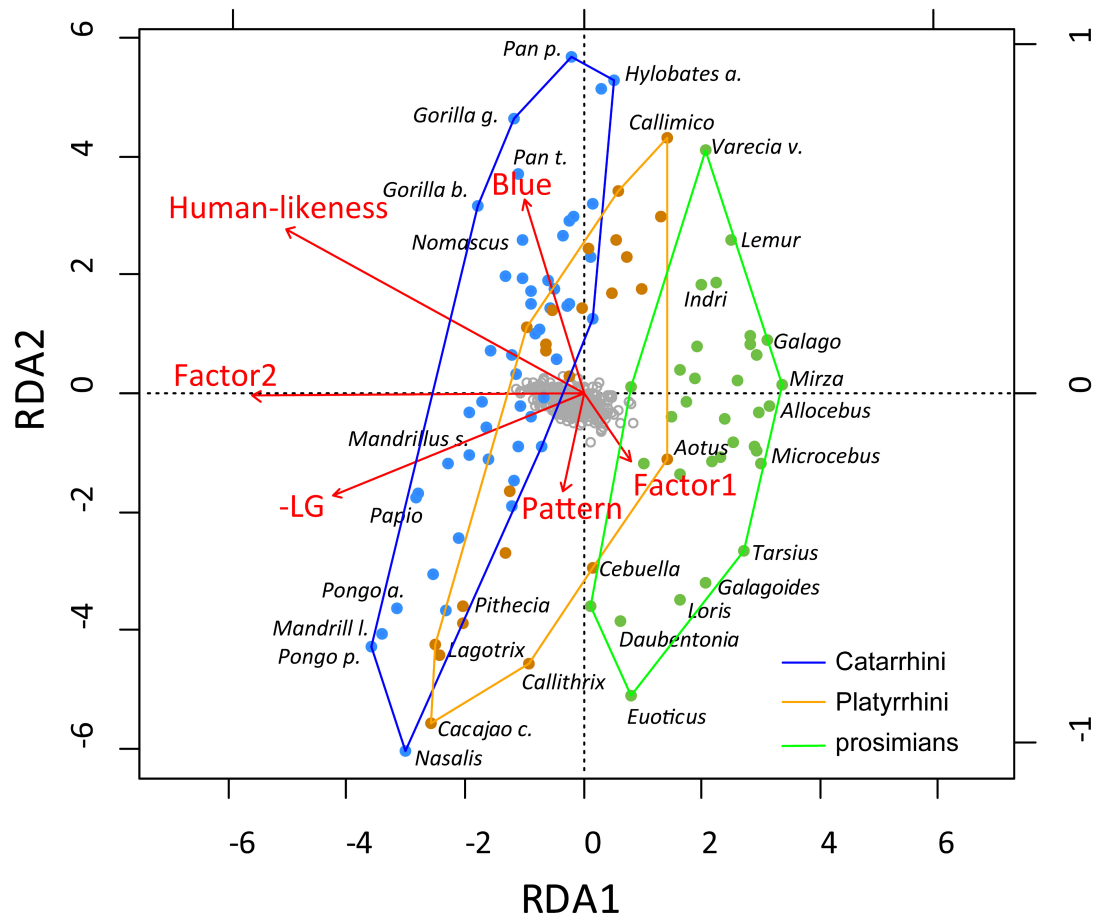


FIGURE 7 | RDA analysis of the factors affecting primate facial attractiveness. RDA1 correlates with Factor2, i.e., the inner facial parts, and RDA2 axis corresponds to the actual attractiveness (see text). The graph shows that the respondents' classification of the primate facial beauty differs among the groups and is mainly based on the inner facial properties of the species.

degree of sexual dimorphism only affects the beauty within the Catarrhines—the male-larger species are perceived as less beautiful. In other words, the respondents rate conspicuous facial features (“extremities”) negatively. Similarly, many researchers agree that distinctive features (caricatures) usually help for better recognition of individual faces (Rhodes et al., 1987; Mauro and Kubovy, 1992; Lee et al., 2000; but see Hancock and Little, 2011), but are rated negatively when being evaluated for attractiveness (Deffenbacher et al., 1998). In other words, conspicuous features are usually rated as unattractive as opposed to average features (e.g., Rhodes and Tremewan, 1996; O’Toole et al., 1999). However, preference for facial averageness may be based on a more general principle as other objects such as dogs, birds, fish, or cars were also found to be preferred when they were of an average shape (Halberstadt and Rhodes, 2000, 2003).

Human-Likeness and the Uncanny Valley Theory

The uncanny valley theory describes an empirical rule first mentioned in an essay by a robotics professor Masahiro Mori in

1970 (and later re-published in English for a broader audience to see; Mori et al., 2012). Mori hypothesized that if a robot resembled humans in appearance, people would feel affinity toward it, up to the point where it was too similar to humans—almost undistinguishable. At that point, people would experience a negative, eerie-like sensation, and he called this descent of affinity the “uncanny valley.” This relationship of human-likeness and attractiveness (in sense of positive affinity toward an object; for a relationship between the emotional ratings of eeriness and attractiveness, see Burleigh et al., 2013) was later tested in a number of papers, which found evidence in support of the uncanny valley theory (e.g., Seyama and Nagayama, 2007; MacDorman et al., 2009; Mitchell et al., 2011). Steckenfinger and Ghazanfar (2009) described a support for this phenomenon even in macaque monkeys, which preferred realistic and stylized macaque faces over faces very close to realism. The reason behind uncanny valley is unclear; often disputed mechanisms include the atypical feature hypothesis or the category conflict hypothesis (Burleigh et al., 2013). In the first case, the effect of uncanny valley is present when evaluating pictures that include an abnormal feature, such as bizarre eyes (Seyama and Nagayama, 2007). In

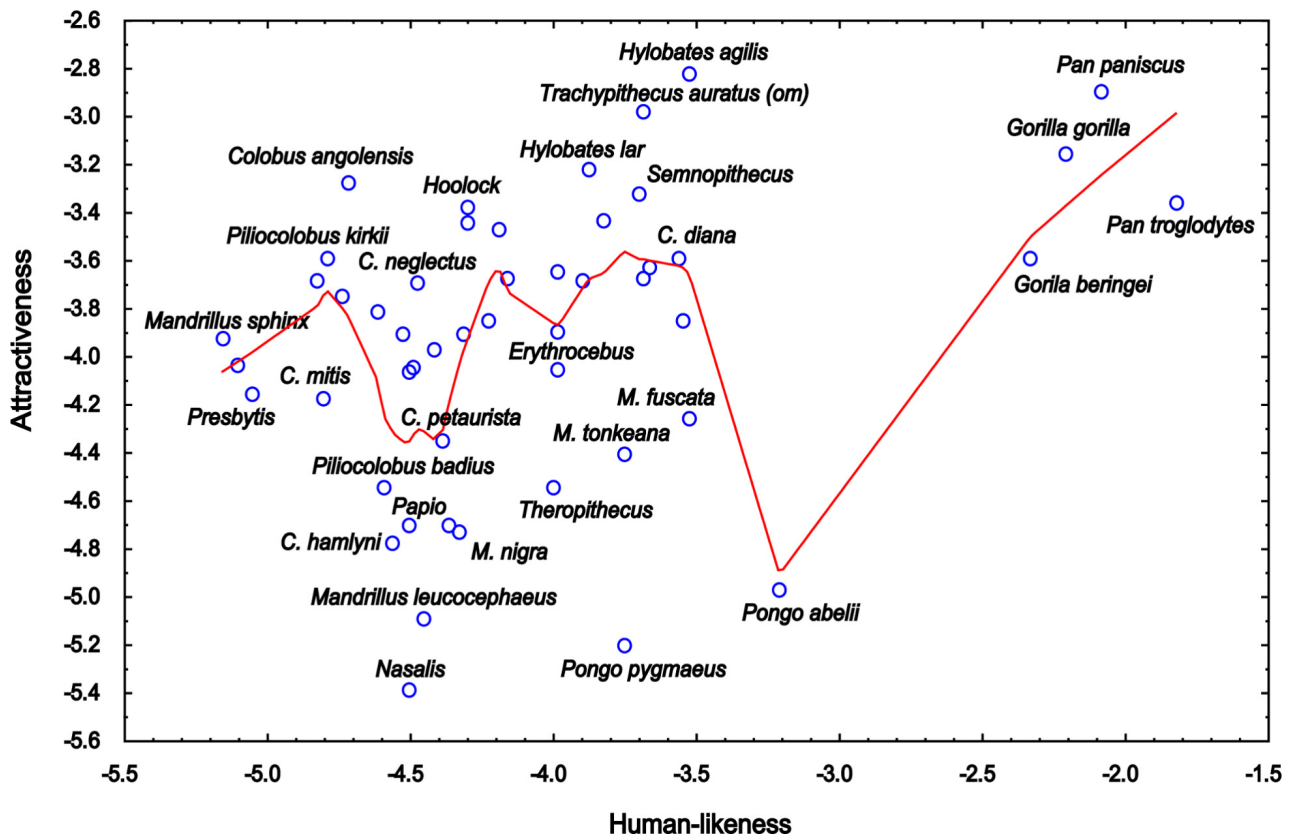


FIGURE 8 | Relationship between the mean rankings of attractiveness of the Catarrhine primate faces and the mean rankings of their human-likeness (the scales were inverted so that the higher value corresponds to higher attractiveness/human-likeness). The LOWESS (Locally Weighted Scatterplot Smoothing) fitting line shows that an “uncanny valley” exists within human-like primates, but not within the ones almost fully human (Gorillas and Chimpanzees). However, many other “overcome the valley” because primates rated as fairly human-like can be rated as either attractive or unattractive depending on their traits (see the text).

the second case, the uncanny valley negatively affects stimuli containing features belonging to multiple categories, eliciting discomfort because it is evaluated as ambiguous and confusing (Saygin et al., 2011). Once the respondents cease to recognize the conflicting object as “human,” the attractiveness returns back to the linear character of the observed attractiveness (or eeriness).

In this matter, our results may seem contradictory as overall the most preferred primates were the prosimians, which were rated as the least human-like. However, within the group of Catarrhine primates, i.e., the group phylogenetically closest to humans (which also includes species most similar to humans, see **Figure 6B**), human-likeness positively affected the rated facial attractiveness. Thus, it is reasonable to examine whether the effect of uncanny valley can be applied to the results within this group. The graph clearly shows that some of the most human-like species do “fall into the valley,” but when looked at in more detail, there are exceptions to this uncanny valley rule. Some primates sharing similar rates of human-likeness fall into the notional valley, while others remain attractive. The unattractive, yet human-like primates, are represented by species such as the Orangutans, Proboscis Monkey, and the Drill, and the reason may be the presence of the abnormal (distinctive) features, as discussed above in Section “SSD, Averageness, and

Facial Extremities”. Interesting fact is that at least some of these features, such as the prolonged nose of the Proboscis monkey, are not perceived as unattractive *per se*: for example, the elephants, elephant shrews, coatis or tapirs all possess a prolonged nose, but they were all rated as very or fairly attractive in a previous study (Frynta et al., 2013). Thus, these results cannot be interpreted simply as a preference for average feature size, but rather as a preference for average feature size when present on a human or human-like animal. Possibly, our complex neural system for facial recognition causes the judgements of “beauty” to be far more strict when judging “humans” (including human-like objects and animals) than different objects (Hanson, 2005). The uncanny valley phenomenon can be in fact linked to the expertise to human faces: the reason why the uncanny effect is so widespread is because every human has an expertise in recognizing humans; however, it is possible that the same effect might be observed in other types of expertise, as these display similar behavioral and neuropsychological pattern (Diamond and Carey, 1986; Xu, 2005; Dufour and Petit, 2010).

Previously, some researchers showed that even distinct, extreme features (up to some point) can be perceived as more attractive than average, if the exaggeration is based on *attractive* features. Such feature then may represent a super-stimulus,

which is a concept derived from ethological studies, describing an object that contains features more accentuated than natural stimulus, which elicits a response more strongly than the stimulus for which it evolved (Tinbergen, 1951). For example, Perrett et al. (1994) report that enhanced female features, including higher cheek bones, thinner jaws, and larger eyes, were rated as more attractive than average (also see Perrett et al., 1998), and similar results were found by Jones and Hill (1993), who found that enlarged proportion of eye width to face height (i.e., feminine/neotenus feature) was preferred more than the average proportions. Could the presence of enhanced attractive features be the reason why in our data, some human-like primates are attractive above the linear relationship between human-likeness and attractiveness (see **Figure 8**)? For example, the Agile Gibbon (*H. agilis*) does have a contrasting, black-and-white face, a feature found to be very attractive in human preferences for animals (Frynta et al., 2013; Lišková et al., 2015). It would be interesting to examine this phenomenon in another study with more controlled, manipulated stimuli.

The support for the uncanny valley theory is ambiguous in our study. Some of the most human-like primates do fall into the notional “valley”; however, some other primates overcome it. Thus, our results rather show that it is the “atypical feature present on a human-like object” that makes some of the primates to be rated very negatively. However, this does not necessarily neglect the uncanny valley. Rather, it supports the atypical feature hypothesis of the uncanny valley theory (Seyama and Nagayama, 2007; MacDorman et al., 2009).

CONCLUSION

In our study, we focused on human evaluation of primate facial attractiveness. We found that there are differences in the evaluation of the three main primate groups. The attractiveness of the Catarrhine primates, i.e., the Old World monkeys, Gibbons, and the Great Apes, was explained by human-likeness, and also by factors similar as those usually utilized when evaluating human facial attractiveness: the inner facial features and SSD (i.e., lack of extreme, conspicuous features). Interestingly, the proportions of inner facial features were only used when evaluating the most human-like primates; in other groups, this factor had no effect, and its importance thus cannot be attributed to the evaluation of faces in general, but only those resembling humans.

The Platyrrhine primates, i.e., the New World monkeys, are phylogenetically more distant to humans. Regarding similarity to humans, they are somewhere between the Catarrhines and the prosimians (see **Figure 6**), and the results explaining their attractiveness scores reflect this. Their attractiveness is

determined by human-likeness, yellowish brown color, and the mean lightness. However, the respondents liked more the monkeys that were scored as less-human like. The orange color, pattern, and SSD are all factors that could not be excluded from the final model but showed to be insignificant. The number of Platyrrhine primates included in the analysis was small though and it is thus possible that a larger sample could reveal significance of these factors. One way or another, the attractiveness of the Old World monkeys seem to be affected by factors that are otherwise reported as affecting evaluation of both human and animal attractiveness. On the contrary, the prosimians were rated as the most beautiful, but our analysis failed to reveal the particular cues responsible for their high scores.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analyzed during the current study are available in the Mendeley repository, <http://dx.doi.org/10.17632/ssv9m953mb.1>.

AUTHOR CONTRIBUTIONS

DF, SR, and EL conceived and designed the research and interpreted the results. SR performed the research and wrote the paper. DF and SR analyzed the data.

FUNDING

This study was a result of the research funded by the project Nr. LO1611, with a financial support from the MEYS under the NPU I program.

ACKNOWLEDGMENTS

We thank Jan Havlíček for a constructive discussion about the topic and Jakub Polák for critical reading of the text and linguistic revision. Moreover, we thank all of the authors who provided us photographs to use in the study, and to all our respondents for their kind participation in the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02439/full#supplementary-material>

REFERENCES

- Akaike, H. (1998). “Information theory and an extension of the maximum likelihood principle,” in *Selected Papers of Hirotugu Akaike*, eds E. Parzen, K. Tanabe, and G. Kitagawa (New York, NY: Springer), 199–213. doi: 10.1007/978-1-4612-1694-0_15
- Archer, J., and Monton, S. (2011). Preferences for infant facial features in pet dogs and cats. *Ethology* 117, 217–226. doi: 10.1111/j.1439-0310.2010.01863.x
- Bonner, L., and Burton, A. M. (2004). 7–11-year-old children show an advantage for matching and recognizing the internal features of familiar faces: evidence against a developmental shift. *Q. J. Exp. Psychol. A* 57, 1019–1029. doi: 10.1080/02724980343000657

- Borgi, M., Cogliati-Dezza, I., Brelsford, V., Meints, K., and Cirulli, F. (2014). Baby schema in human and animal faces induces cuteness perception and gaze allocation in children. *Front. Psychol.* 5:411. doi: 10.3389/fpsyg.2014.00411
- Borkenau, P., Brecke, S., Möttig, C., and Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *J. Res. Pers.* 43, 703–706. doi: 10.1016/j.jrp.2009.03.007
- Burleigh, T. J., Schoenherr, J. R., and Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Comput. Hum. Behav.* 29, 759–771. doi: 10.1016/j.chb.2012.11.021
- Calvo, M. G., and Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: an integrative review. *Cogn. Emot.* 30, 1081–1106. doi: 10.1080/02699931.2015.1049124
- Campbell, R., Walker, J., and Baron-Cohen, S. (1995). The development of differential use of inner and outer face features in familiar face identification. *J. Exp. Child Psychol.* 59, 196–210. doi: 10.1006/jecp.1995.1009
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* 6, 284–290. doi: 10.1037/1040-3590.6.4.284
- Coetzee, V., Faerber, S. J., Greeff, J. M., Lefevre, C. E., Re, D. E., and Perrett, D. I. (2012). African perceptions of female attractiveness. *PLoS One* 7:e48116. doi: 10.1371/journal.pone.0048116
- DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *P. Roy. Soc. Lond. B* 272, 919–922. doi: 10.1098/rspb.2004.3003
- Deffenbacher, K. A., Vetter, T., Johanson, J., and O'Toole, A. J. (1998). Facial aging, attractiveness, and distinctiveness. *Perception* 27, 1233–1243. doi: 10.1068/p271233
- Diamond, R., and Carey, S. (1986). Why faces are and are not special: an effect of expertise. *J. Exp. Psychol. Gen.* 115, 107–117. doi: 10.1037/0096-3445.115.2.107
- Dixon, A. F. (2012). *Comparative Studies of the Prosimians, Monkeys, Apes, and Humans*. Oxford: Oxford University Press, 808. doi: 10.1093/acprof:osobl/9780199544646.001.0001
- Dufour, V., and Petit, O. (2010). Recognition of monkey faces by monkey experts. *J. Ethol.* 28, 231–238. doi: 10.1007/s10164-009-0174-8
- Ekman, P., and Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motiv. Emot.* 10, 159–168. doi: 10.1007/BF00992253
- Elia, I. E. (2013). A foxy view of human beauty: implications of the farm fox experiment for understanding the origins of structural and experiential aspects of facial attractiveness. *Q. Rev. Biol.* 88, 163–183. doi: 10.1086/671486
- Elliot, A. J., Niesta Kayser, D., Greitemeyer, T., Lichtenfeld, S., Gramzow, R. H., Maier, M. A., et al. (2010). Red, rank, and romance in women viewing men. *J. Exp. Psychol. Gen.* 139, 399–417. doi: 10.1037/a0019689
- Ellis, H. D., Shepherd, J. W., and Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception* 8, 431–439. doi: 10.1068/p080431
- Fink, B., Grammer, K., and Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evol. Hum. Behav.* 27, 433–442. doi: 10.1016/j.evolhumbehav.2006.08.007
- Fink, B., Grammer, K., and Thornhill, R. (2001). Human (*Homo sapiens*) facial attractiveness in relation to skin texture and color. *J. Comp. Psychol.* 115, 92–99. doi: 10.1037/0735-7036.115.1.92
- Fink, B., Matts, P. J., D'Emiliano, D., Bunse, L., Weege, B., and Röder, S. (2012). Colour homogeneity and visual perception of age, health and attractiveness of male facial skin. *J. EADV* 26, 1486–1492. doi: 10.1111/j.1468-3083.2011.04316.x
- Fink, B., and Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Curr. Dir. Psychol. Sci.* 11, 154–158. doi: 10.1111/1467-8721.00190
- Finstermeier, K., Zinner, D., Brameier, M., Meyer, M., Kreuz, E., Hofreiter, M., et al. (2013). A mitogenomic phylogeny of living primates. *PLoS One* 8:e69504. doi: 10.1371/journal.pone.0069504
- Fisher, M. L. (2004). Female intrasexual competition decreases female facial attractiveness. *P. Roy. Soc. Lond. B* 271, S283–S285. doi: 10.1098/rsbl.2004.0160
- Fridlund, A. J. (1994). *Human Facial Expression: An Evolutionary View*. San Diego, CA: Academic Press.
- Frisby, C. M. (2006). “Shades of beauty”: examining the relationship of skin color to perceptions of physical attractiveness. *Facial Plast. Sur.* 22, 175–179. doi: 10.1055/s-2006-950174
- Frynta, D., Landová, E., and Lišková, S. (2014). “Animal beauty, cross-cultural perceptions,” in *Encyclopedia of Quality of Life and Well-Being Research*, ed. A. C. Michalos (Dordrecht: Springer).
- Frynta, D., Lišková, S., Bültmann, S., and Burda, H. (2010). Being attractive brings advantages: the case of parrot species in captivity. *PLoS One* 5:e12568. doi: 10.1371/journal.pone.0012568
- Frynta, D., Marešová, E., Landová, E., Lišková, S., Šimková, O., Tichá, I., et al. (2009). “Are animals in zoos rather conspicuous than endangered?,” in *Endangered Species - New Research*, eds A. M. Columbus and L. Kuznetsov (New York, NY: Nova Science Publishers, Inc), 299–341.
- Frynta, D., Šimková, O., Lišková, S., and Landová, E. (2013). Mammalian collection on Noah's ark: the effects of beauty, brain and body size. *PLoS One* 8:e63110. doi: 10.1371/journal.pone.0063110
- Ge, L., Anzures, G., Wang, Z., Kelly, D. J., Pascalis, O., Quinn, P. C., et al. (2008). An inner face advantage in children's recognition of familiar peers. *J. Exp. Child Psychol.* 101, 124–136. doi: 10.1016/j.jecp.2008.05.006
- Ge, L., Zhang, H., Wang, Z., Quinn, P. C., Pascalis, O., Kelly, D., et al. (2009). Two faces of the other-race effect: recognition and categorisation of Caucasian and Chinese faces. *Perception* 38, 1199–1210. doi: 10.1068/p6136
- Geissmann, T., Groves, C. P., and Roos, C. (2004). The *Tenasserim lutung*, *Trachypitecus barbei* (Blyth, 1847) (Primates: Cercopithecidae): description of a live specimen, and a reassessment of phylogenetic affinities, taxonomic history, and distribution. *Bijdr. Dierkunde* 73, 271–282.
- Gordon, A. D. (2006). Scaling of size and dimorphism in primates II: macroevolution. *Int. J. Primatol.* 27, 63–105. doi: 10.1007/s10764-005-9004-1
- Gothard, K. M., Brooks, K. N., and Peterson, M. A. (2009). Multiple perceptual strategies used by macaque monkeys for face recognition. *Anim. Cogn.* 12, 155–167. doi: 10.1007/s10071-008-0179-7
- Grammer, K., and Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness. *J. Comp. Psychol.* 108, 233–242. doi: 10.1037/0735-7036.108.3.233
- Halberstadt, J., and Rhodes, G. (2000). The attractiveness of nonface averages: implications for an evolutionary explanation of the attractiveness of average faces. *Psychol. Sci.* 11, 285–289. doi: 10.1111/1467-9280.00257
- Halberstadt, J., and Rhodes, G. (2003). It's not just average faces that are attractive: computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychon. B. Rev.* 10, 149–156. doi: 10.3758/BF03196479
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor. Quant. Methods Psychol.* 8, 23–34. doi: 10.20982/tqmp.08.1.p023
- Hancock, P. J., and Little, A. C. (2011). Adaptation may cause some of the face caricature effect. *Perception* 40, 317–322. doi: 10.1068/p6865
- Hanson, D. (2005). “Expanding the aesthetic possibilities for humanoid robots,” in *IEEE-RAS International Conference on Humanoid Robots*, (Birmingham), 24–31.
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0
- Hecht, J., and Horowitz, A. (2015). Seeing dogs: human preferences for dog physical attributes. *Anthrozoös* 28, 153–163. doi: 10.2752/089279315X14129350722217
- Hughes, K. D., Higham, J. P., Allen, W. L., Elliot, A. J., and Hayden, B. Y. (2015). Extraneous color affects female macaques' gaze preference for photographs of male conspecifics. *Evol. Hum. Behav.* 36, 25–31. doi: 10.1016/j.evolhumbehav.2014.08.003
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L., Feinberg, D. R., et al. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *P. Roy. Soc. Lond. B* 277, 617–624. doi: 10.1098/rspb.2009.1575
- Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception* 33, 569–576. doi: 10.1068/p3463
- Jones, D., and Hill, K. (1993). Criteria of facial attractiveness in five populations. *Hum. Nature* 4, 271–296. doi: 10.1007/BF02692202

- Kay, R. F., Plavcan, J. M., Glander, K. E., and Wright, P. C. (1988). Sexual selection and canine dimorphism in New World monkeys. *Am. J. Phys. Anthropol.* 77, 385–397. doi: 10.1002/ajpa.1330770311
- Komori, M., Kawamura, S., and Ishihara, S. (2009). Averageness or symmetry: which is more important for facial attractiveness? *Acta Psychol.* 131, 136–142. doi: 10.1016/j.actpsy.2009.03.008
- Kościński, K. (2007). Facial attractiveness: general patterns of facial preferences. *Anthropol. Rev.* 70, 45–79. doi: 10.2478/v10044-008-0001-9
- Landová, E., Bakhshaliyeva, N., Janovcová, M., Peléšková, Š., Suleymanova, M., Polák, J., et al. (2018). Association between fear and beauty evaluation of snakes: cross-cultural findings. *Front. Psychol.* 9:333. doi: 10.3389/fpsyg.2018.00333
- Landová, E., Marešová, J., Šimková, O., Cikánová, V., and Frynta, D. (2012). Human responses to live snakes and their photographs: evaluation of beauty and fear of the king snakes. *J. Environ. Psychol.* 32, 69–77. doi: 10.1016/j.jenvp.2011.10.005
- Lee, K., Byatt, G., and Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychol. Sci.* 11, 379–385. doi: 10.1111/1467-9280.00274
- Lei, R. (2008). *Nocturnal Lemur Diversity at Masoala National Park* (No. 53). Lubbock, TX: Museum of Texas Tech University. doi: 10.5962/bhl.title.142673
- Leutenegger, W., and Kelly, J. T. (1977). Relationship of sexual dimorphism in canine size and body size to social, behavioral, and ecological correlates in anthropoid primates. *Primates* 18, 117–136. doi: 10.1007/BF02382954
- Lindénfors, P., and Tullberg, B. S. (1998). Phylogenetic analyses of primate size evolution: the consequences of sexual selection. *Biol. J. Linn. Soc.* 64, 413–447. doi: 10.1111/j.1095-8312.1998.tb00342.x
- Lišková, S., and Frynta, D. (2013). What determines bird beauty in human eyes? *Anthrozoös* 26, 27–41. doi: 10.2752/175303713X13534238631399
- Lišková, S., Landová, E., and Frynta, D. (2015). Human preferences for colorful birds: vivid colors or pattern? *Evol. Psychol.* 13, 339–359. doi: 10.1177/147470491501300203
- Little, A. C., and Hancock, P. J. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *Brit. J. Psychol.* 93, 451–464. doi: 10.1348/000712602761381349
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. T. Roy. Soc. B* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Little, A. C., Jones, B. C., DeBruine, L. M., and Feinberg, D. R. (2008). Symmetry and sexual dimorphism in human faces: interrelated preferences suggest both signal quality. *Behav. Ecol.* 19, 902–908. doi: 10.1093/beheco/arn049
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *P. Roy. Soc. Lond. B* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Lovich, J. E., and Gibbons, J. W. (1992). A review of techniques for quantifying sexual size dimorphism. *Growth Dev. Aging* 56, 269–281.
- MacDorman, K. F., Green, R. D., Ho, C. C., and Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Comput. Hum. Behav.* 25, 695–710. doi: 10.1016/j.chb.2008.12.026
- Marešová, J., Landová, E., and Frynta, D. (2009). What makes some species of milk snakes more attractive to humans than others? *Theor. Biosci.* 128, 227–235. doi: 10.1007/s12064-009-0075-y
- Matts, P. J., Fink, B., Grammer, K., and Burquest, M. (2007). Color homogeneity and visual perception of age, health, and attractiveness of female facial skin. *J. Am. Acad. Dermatol.* 57, 977–984. doi: 10.1016/j.jaad.2007.07.040
- Maurer, D., Le Grand, R., and Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn. Sci.* 6, 255–260. doi: 10.1016/S1364-6613(02)01903-4
- Mauro, R., and Kubovy, M. (1992). Caricature and face recognition. *Mem. Cogn.* 20, 433–440. doi: 10.3758/BF03210927
- Mayor, M. I., Sommer, J. A., Houck, M. L., Zaonarivelo, J. R., Wright, P. C., Ingram, C., et al. (2004). Specific status of *Propithecus* spp. *Int. J. Primatol.* 25, 875–900. doi: 10.1023/B:IJOP.0000029127.31190.e9
- McGraw, K. O., and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* 1, 30–46. doi: 10.1037/1082-989X.1.1.30
- Michel, C., Caldara, R., and Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Vis. Cogn.* 14, 55–73. doi: 10.1080/13506280500158761
- Mitchell, W. J., Szerszen, K. A. Sr., Lu, A. S., Schermerhorn, P. W., Scheutz, M., and MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *IPerception* 2, 10–12. doi: 10.1068/i0415
- Mittermeier, R. A., Louis, E. E., Richardson, M., Schwitzer, C., Langrand, O., Rylands, A. B., et al. (2010). *Lemurs of Madagascar*. Illustrated by S. D. Nash, 3rd Edn. Arlington, VA: Conservation International.
- Mitteroecker, P., Gunz, P., Windhager, S., and Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix It. J. Mamm.* 24, 59–66. doi: 10.4404/hystrix-24.1-6369
- Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robot Autom. Mag.* 19, 98–100. doi: 10.1109/MRA.2012.2192811
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2017). *Vegan: Community Ecology Package. R Package Version 2.4-5*.
- O'Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., and Blanz, V. (1999). 3D shape and 2D surface textures of human faces: the role of “averages” in attractiveness and age. *Image Vis. Comput.* 18, 9–19. doi: 10.1016/S0262-8856(99)00012-8
- Pascalis, O., and Bachevalier, J. (1998). Face recognition in primates: a cross-species study. *Behav. Process.* 43, 87–96. doi: 10.1016/S0376-6357(97)00090-9
- Pastorini, J., Forstner, M. R., and Martin, R. D. (2001). Phylogenetic history of sifakas (*Propithecus*: lemuriformes) derived from mtDNA sequences. *Am. J. Primatol.* 53, 1–17. doi: 10.1002/1098-2345(200101)53:1<1::AID-AJPI>3.0.CO;2-J
- Penton-Voak, I. S., Pound, N., Little, A. C., and Perrett, D. I. (2006). Personality judgments from natural and composite facial images: more evidence for a “kernel of truth” in social perception. *Soc. Cogn.* 24, 607–640. doi: 10.1521/soco.2006.24.5.607
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., and Edwards, R. (1999). Symmetry and human facial attractiveness. *Evol. Hum. Behav.* 20, 295–307. doi: 10.1016/S1090-5138(99)00014-8
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Perrett, D. I., May, K. A., and Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature* 368, 239–242. doi: 10.1038/368239a0
- Pothos, E. M., and Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cogn. Sci.* 26, 303–343. doi: 10.1016/S0364-0213(02)00064-2
- Pothos, E. M., and Close, J. (2008). One or two dimensions in spontaneous classification: a simplicity approach. *Cognition* 107, 581–602. doi: 10.1016/j.cognition.2007.11.007
- Ptáčeková, J., Landová, E., Lišková, S., Kubina, A., and Frynta, D. (2017). Are the aesthetic preferences towards snake species already formed in pre-school aged children? *Eur. J. Dev. Psychol.* 14, 16–31. doi: 10.1080/17405629.2016.1144507
- Purvis, A. (1995). A composite estimate of primate phylogeny. *Philos. T. Roy. Soc. B* 348, 405–421. doi: 10.1098/rstb.1995.0078
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, Austria.
- Rádlová, S., Viktorin, P., and Frynta, D. (2016). *Barvocuc 2.0, Software for Color Image Analysis*. Available at: http://www.nudz.cz/en/w_group/wg-on-emotions-and-phobias-triggered-by-animals/
- Rankin, M., and Borah, G. L. (2003). Perceived functional impact of abnormal facial appearance. *Plast. Reconstr. Surg.* 111, 2140–2146. doi: 10.1097/01.PRS.0000060105.63335.0C
- Re, D. E., Whitehead, R. D., Xiao, D., and Perrett, D. I. (2011). Oxygenated-blood colour change thresholds for perceived facial redness, health, and attractiveness. *PLoS One* 6:e17859. doi: 10.1371/journal.pone.0017859
- Rhodes, G., Brennan, S., and Carey, S. (1987). Identification and ratings of caricatures: implications for mental representations of faces. *Cogn. Psychol.* 19, 473–497. doi: 10.1016/0010-0285(87)90016-8
- Rhodes, G., Hickford, C., and Jeffery, L. (2000). Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive? *Brit. J. Psychol.* 91, 125–140. doi: 10.1348/000712600161718

- Rhodes, G., and Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychol. Sci.* 7, 105–110. doi: 10.1111/j.1467-9280.1996.tb00338.x
- Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., and Akamatsu, S. (2001). Attractiveness of facial averageness and symmetry in non-Western cultures: in search of biologically based standards of beauty. *Perception* 30, 611–625. doi: 10.1068/p3123
- Rowe, N. (1996). *The Pictorial Guide to the Living Primates*. East Hampton, NY: Pogonias Press, 263.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychol. Bull.* 115, 102–141. doi: 10.1037/0033-2909.115.1.102
- Samson, N., Fink, B., and Matts, P. J. (2010). Visible skin condition and perception of human facial appearance. *Int. J. Cosmetic Sci.* 32, 167–184. doi: 10.1111/j.1468-2494.2009.00535.x
- Santana, S. E., Alfaro, J. L., and Alfaro, M. E. (2012). Adaptive evolution of facial colour patterns in Neotropical primates. *P. Roy. Soc. Lond. B* 279, 2204–2211. doi: 10.1098/rspb.2011.2326
- Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., and Frith, C. (2011). The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Soc. Cogn. Affect. Neurosci.* 7, 413–422. doi: 10.1093/scan/nsr025
- Scheib, J. E., Gangestad, S. W., and Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *P. Roy. Soc. Lond. B* 266, 1913–1917. doi: 10.1098/rspb.1999.0866
- Seyama, J. I., and Nagayama, R. S. (2007). The uncanny valley: effect of realism on the impression of artificial human faces. *Presence Teleop. Virt.* 16, 337–351. doi: 10.1162/pres.16.4.337
- Sforza, C., Laino, A., D'Alessio, R., Grandi, G., Catti, F., and Ferrario, V. F. (2007). Three-dimensional facial morphometry of attractive adolescent boys and girls. *Prog. Orthod.* 8, 268–281.
- Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86:420. doi: 10.1037/0033-2909.86.2.420
- Sobel, I. (1978). Neighborhood coding of binary images for fast contour following and general binary array processing. *Comput. Vis. Graph.* 8, 127–135. doi: 10.1016/S0146-664X(78)80020-3
- Stadie, W. C. (1919). The oxygen of the arterial and venous blood in pneumonia and its relation to cyanosis. *J. Exp. Med.* 30, 215–240. doi: 10.1084/jem.30.3.215
- StatSoft (2010). *STATISTICA (Data Analysis Software System), Version 9.1*. Available at: www.statsoft.com
- Steckenfinger, S. A., and Ghazanfar, A. A. (2009). Monkey visual behavior falls into the uncanny valley. *P. Natl. Acad. Sci.* 106, 18362–18366. doi: 10.1073/pnas.0910063106
- Stephen, I. D., Scott, I. M., Coetzee, V., Pound, N., Perrett, D. I., and Penton-Voak, I. S. (2012). Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol. Hum. Behav.* 33, 260–267. doi: 10.1016/j.evolhumbehav.2011.10.003
- Stephen, I. D., Smith, M. J. L., Stirrat, M. R., and Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *Int. J. Primatol.* 30, 845–857. doi: 10.1007/s10764-009-9380-z
- Stirrat, M., and Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychol. Sci.* 21, 349–354. doi: 10.1177/0956797610362647
- Tanaka, J. W., Kiefer, M., and Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition* 93, B1–B9. doi: 10.1016/j.cognition.2003.09.011
- Taubert, J. (2009). Chimpanzee faces are 'special' to humans. *Perception* 38, 343–356. doi: 10.1068/p6254
- Thomaz, C. E. (2012). *FEI Face Database*. Available at: <http://fei.edu.br/~cet/facedatabase.html>
- Thornhill, R., and Gangestad, S. W. (1999). Facial attractiveness. *Trends Cogn. Sci.* 3, 452–460. doi: 10.1016/S1364-6613(99)01403-5
- Tinbergen, N. (1951). *The Study of Instinct*. New York, NY: Oxford Univ. Press, 228.
- Tobiasen, J. M. (1987). Social judgments of facial deformity. *Cleft Palate J.* 24, 323–327.
- Trujillo, L. T., Jankowitsch, J. M., and Langlois, J. H. (2014). Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cogn. Affect. Behav. Neurosci.* 14, 1061–1076. doi: 10.3758/s13415-013-0230-2
- Turati, C., Macchi Cassia, V., Simion, F., and Leo, I. (2006). Newborns' face recognition: role of inner and outer facial features. *Child Dev.* 77, 297–311. doi: 10.1111/j.1467-8624.2006.00871.x
- Valenzano, D. R., Mennucci, A., Tartarelli, G., and Cellerino, A. (2006). Shape analysis of female facial attractiveness. *Vis. Res.* 46, 1282–1291. doi: 10.1016/j.visres.2005.10.024
- Van den Berghe, P. L., and Frost, P. (1986). Skin color preference, sexual dimorphism and sexual selection: a case of gene culture co-evolution? *Ethnic Racial Stud.* 9, 87–113. doi: 10.1080/01419870.1986.9993516
- Waite, C., Gerald, M. S., Little, A. C., and Kraisselburd, E. (2006). Selective attention toward female secondary sexual color in male rhesus macaques. *Am. J. Primatol.* 68, 738–744. doi: 10.1002/ajp.20264
- Waite, C., Little, A. C., Wolfensohn, S., Honess, P., Brown, A. P., Buchanan-Smith, H. M., et al. (2003). Evidence from rhesus macaques suggests that male coloration plays a role in female primate mate choice. *P. Roy. Soc. Lond. B* 270, S144–S146. doi: 10.1098/rsbl.2003.0065
- Want, S. C., Pascalis, O., Coleman, M., and Blades, M. (2003). Recognizing people from the inner or outer parts of their faces: developmental data concerning 'unfamiliar' faces. *Brit. J. Dev. Psychol.* 21, 125–135. doi: 10.1348/026151003321164663
- Weston, E. M., Friday, A. E., Johnstone, R. A., and Schrenk, F. (2004). Wide faces or large canines? The attractive versus the aggressive primate. *P. Roy. Soc. Lond. B* 271, S416–S419. doi: 10.1098/rsbl.2004.0203
- Wilson, D. E., and Reeder, D. M. (eds) (2005). *Mammal Species of the World. A Taxonomic and Geographic Reference*, 3rd Edn. Baltimore, MD: Johns Hopkins University Press.
- Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cereb. Cortex* 15, 1234–1242. doi: 10.1093/cercor/bhi006
- Yoder, A. D., and Irwin, J. A. (1999). Phylogeny of the Lemnidae: effects of character and taxon sampling on resolution of species relationships within Eulemur. *Cladistics* 15, 351–361. doi: 10.1111/j.1096-0031.1999.tb00271.x
- Yovel, G. (2016). Neural and cognitive face-selective markers: an integrative review. *Neuropsychologia* 83, 5–13. doi: 10.1016/j.neuropsychologia.2015.09.026
- Zebrowitz, L. A., Hall, J. A., Murphy, N. A., and Rhodes, G. (2002). Looking smart and looking good: facial cues to intelligence and their origins. *Pers. Soc. Psychol. B* 28, 238–249. doi: 10.1177/0146167202282009
- Zelditch, M. L., Swiderski, D. L., and Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Cambridge, MA: Academic Press.
- Zhang, D., Zhao, Q., and Chen, F. (2011). Quantitative analysis of human facial beauty using geometric features. *Pattern Recogn.* 44, 940–950. doi: 10.1016/j.patcog.2010.10.013
- Zhao, Y., Zhen, Z., Liu, X., Song, Y., and Liu, J. (2018). The neural network for face recognition: insights from an fMRI study on developmental prosopagnosia. *Neuroimage* 169, 151–161. doi: 10.1016/j.neuroimage.2017.12.023

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Rádlová, Landová and Frynta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Facial Adiposity, Attractiveness, and Health: A Review

Stefan de Jager^{1,2*}, Nicoleen Coetzee² and Vinet Coetzee³

¹ Department of Psychology, Sefako Makgatho Health Sciences University, Ga-Rankuwa, South Africa, ² Department of Psychology, University of Pretoria, Pretoria, South Africa, ³ Department of Genetics, Biochemistry and Microbiology, University of Pretoria, Pretoria, South Africa

OPEN ACCESS

Edited by:

Kok Wei Tan,
University of Reading Malaysia,
Malaysia

Reviewed by:

Barnaby James Wyld Dixon,
The University of Queensland,
Australia

Danielle Leigh Wagstaff,
Federation University, Australia

Shen Liu,
University of Science and Technology
of China, China

*Correspondence:

Stefan de Jager
gsdejager@gmail.com

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 September 2018

Accepted: 29 November 2018

Published: 21 December 2018

Citation:

de Jager S, Coetzee N and Coetzee V
(2018) Facial Adiposity, Attractiveness,
and Health: A Review.
Front. Psychol. 9:2562.
doi: 10.3389/fpsyg.2018.02562

The relationship between facial cues and perceptions of health and attractiveness in others plays an influential role in our social interactions and mating behaviors. Several facial cues have historically been investigated in this regard, with facial adiposity being the newest addition. Evidence is mounting that a robust link exists between facial adiposity and attractiveness, as well as perceived health. Facial adiposity has also been linked to various health outcomes such as cardiovascular disease, respiratory disease, blood pressure, immune function, diabetes, arthritis, oxidative stress, hormones, and mental health. Though recent advances in the analysis of facial morphology has led to significant strides in the description and quantification of facial cues, it is becoming increasingly clear that there is a great deal of nuance in the way that humans use and integrate facial cues to form coherent social or health judgments of others. This paper serves as a review of the current literature on the relationship between facial adiposity, attractiveness, and health. A key component in utilizing facial adiposity as a cue to health and attractiveness perceptions is that people need to be able to estimate body mass from facial cues. To estimate the strength of the relationship between perceived facial adiposity and body mass, a meta-analysis was conducted on studies that quantified the relationship between perceived facial adiposity and BMI/percentage body fat. Summary effect size estimates indicate that participants could reliably estimate BMI from facial cues alone ($r = 0.71$, $n = 458$).

Keywords: facial adiposity, attractiveness, perceived health, health outcomes, BMI, percentage body fat, meta-analysis

INTRODUCTION

Facial appearance in humans conveys a substantial amount of non-verbal information when it comes to our interactions with others. These include judgments of health (Rhodes et al., 2007; Coetzee et al., 2009; Stephen et al., 2009; Phalane et al., 2017), attractiveness (Rhodes, 2006; Coetzee et al., 2012; Foo et al., 2017), leadership ability (Little et al., 2007; Re and Perrett, 2014) and even academic ability (Zebrowitz et al., 2002; Talamas et al., 2016), to name but a few. Naturally, the evolutionary, social and behavioral implications of the way in which we perceive and react to other people's faces, has generated significant interest in the scientific community. Several facial cues have been identified as integral aspects of our judgments of others, including sexual dimorphism (Perrett et al., 1998; Thornhill and Gangestad, 2006), symmetry (Perrett et al., 1999; Scheib et al., 1999; Jones et al., 2001), averageness (Grammer and Thornhill, 1994; Rhodes et al., 2001), skin condition (Jones et al., 2004; Stephen et al., 2011), and facial adiposity, or perceived weight in the face (Coetzee et al., 2009; Tinlin et al., 2013). Of particular interest to scientists is the relationship between facial cues and perceptions of health and attractiveness

in others, as these play a particularly influential role in our social interactions and mating behaviors. Since mate choice plays a central role in evolutionary psychology, researchers are eager to understand the relative contribution that various facial cues make in shaping our perceptions and judgments of health and attractiveness (Rhodes, 2006).

Attractiveness plays a prominent role in our everyday interactions with others and research has consistently shown that attractive people are judged more positively in general compared to unattractive people (see Rhodes, 2006; Little et al., 2011). Attractive people are also typically judged to be healthier compared to unattractive people (Kalick et al., 1998; Shackelford and Larsen, 1999; Boothroyd et al., 2013). One potential reason for this is that facial cues associated with attractiveness could also be linked to various health markers. To this end, humans evolved to be particularly attentive to facial cues associated with attractiveness, as these cues allow us to make inferences about other people's health and genetic fitness, or even behavioral tendencies (Rhodes et al., 2001). For example, fitness-related theories of human behavior suggest that key phenotypic cues influence our judgments of others because they evolved as cues to general health and mate quality (Langlois et al., 2000). One example of such a fitness-related theory is the "good genes" hypothesis, which postulates that female mate choice is heavily influenced by phenotypic cues, since they aid females in making snap judgments about men's current health, genetic quality and fertility, and thus suitability for mating (Hamilton and Zuk, 1982; Thornhill and Gangestad, 1993, 1999).

Choosing the right mate could provide substantial direct and indirect benefits to females and their offspring. For example, according to the immunocompetence handicap hypothesis (Folstad and Karter, 1992) males that display strong secondary male sexual traits are perceived as more attractive by females, as these secondary sexual traits may potentially serve as good indicators of immune function in males. The reason for this is that the testosterone that is responsible for the development of these secondary male sexual traits also has an immunosuppressive effect, and only good quality males can therefore afford to display them. By choosing a male with a strong immune system as a mating partner, a female can simultaneously lessen the risk of pathogen exposure, while also maximizing the robustness of her offspring in terms of immune function. It should be mentioned, however, that some studies have found very little to no evidence for a relationship between testosterone and immune suppression in mammals and humans, respectively, thus casting at least some doubt on the viability of the immunocompetence handicap hypothesis (Roberts et al., 2004; Nowak et al., 2018).

From an evolutionary point of view, preferences for certain facial cues as reliable indicators of health and genetic quality, can only be a coherent hypothesis if two conditions are met: (i) there should be agreement between people regarding the facial features that they find attractive (Langlois et al., 2000; Rhodes, 2006) and (ii) perceptions of attractiveness and health should be related to actual health outcomes or genetic quality (Coetzee et al., 2009; Rantala et al., 2013a). At present there is at least some empirical support for the notion that people can reliably detect and agree on facial cues that contribute to facial attractiveness and perceptions

of health (Perrett et al., 1998; Langlois et al., 2000; Rhodes et al., 2001; Re and Rule, 2016a). This statement does come with a caveat though, as numerous studies have demonstrated that preferences for certain facial cues, such as facial adiposity for example, are likely mediated by environmental and cultural factors, including resource scarcity (Batres and Perrett, 2014, 2017), exposure to media beauty ideals (Batres and Perrett, 2014), and own ethnicity familiarity (Coetzee et al., 2014; Batres et al., 2017).

While there is strong support for the link between various facial cues and attractiveness, the link between facial attractiveness and actual health outcomes has been mixed (Weeden and Sabini, 2005; Rhodes et al., 2007; Foo et al., 2017; Cai et al., in press). For example, Kalick et al. (1998) found no reliable relationship between rated facial attractiveness and general health rating by physicians for people in adolescence, middle adulthood, or late adulthood. The researchers did, however, find that there was a correlation between rated facial attractiveness and perceived health for both males and females. After controlling for the mediating relationship that attractiveness may play between perceived health and actual health outcomes, a statistically significant correlation between perceived health and actual health was found. The researchers conclude that, paradoxically, attractiveness can cause a halo effect and serve to suppress accurate judgments of health, instead of enhancing them. In partial contrast to Kalick et al. (1998), Henderson and Anglin (2003) did find a significant correlation between rated attractiveness and longevity for 50 facial photographs taken from a high school yearbook, although it is not entirely clear how comparable longevity (Henderson and Anglin, 2003) and general health measures (Kalick et al., 1998) are within this context. For a review on the relationship between attractiveness and health see Weeden and Sabini (2005) and Rhodes (2006).

The relationship between individual facial cues associated with attractiveness and actual health outcomes has also been met with mixed results. For example, a study by Rhodes et al. (2001), using a very similar dataset to Kalick et al. (1998), found a relationship between facial averageness and perceived health, as well as facial symmetry and perceived health for both males and females. However, when the researchers investigated the potential relationship between actual health outcomes and facial averageness, a correlation was found only for male childhood health ($r = 0.28$), female adolescent health ($r = 0.14$) and the current health for females ($r = 0.25$). For rated asymmetry, no link was found between health outcomes and perceived symmetry of the participant's faces. One potential explanation for this inconsistent trend is that general health is a very broad concept that probably reflects a wide variety of health markers, which may, or may not, be linked to particular facial cues or even general facial attractiveness.

Facial Adiposity as a Cue to Perceived Health and Attractiveness

One of the facial cues that has consistently been associated with both attractiveness (Coetzee, 2011; Coetzee et al., 2012; Rantala et al., 2013a; Foo et al., 2017; Phalane et al., 2017),

perceived health (Coetzee et al., 2009; Fisher et al., 2014b; Han et al., 2016; Windhager et al., 2018) and actual health outcomes (Coetzee et al., 2009; Reither et al., 2009; Rantala et al., 2013a; Tinlin et al., 2013; Martinson and Vasunilashorn, 2016) is facial adiposity. Facial adiposity, or perceived weight in the face, was demonstrated to be a reliable cue to health in a study published by Coetzee et al. (2009) where higher facial adiposity ratings were linked to perceived health and attractiveness, as well as increased risk of infection and cardiovascular-illness. Studies also consistently report that people with lower perceived facial adiposity are rated as more attractive and healthier compared to people with higher perceived facial adiposity (Klaczynski et al., 2009; Rantala et al., 2013a; Han et al., 2016; Foo et al., 2017). A study by Foo et al. (2017) showed that facial adiposity was a better predictor of attractiveness compared to sexual dimorphism, averageness, and symmetry, for male faces. The researchers also found that, for females faces, facial adiposity squared and sexual dimorphism were the best predictors of female facial attractiveness, while facial adiposity was also the strongest predictor of perceived health for male faces, while sexual dimorphism was the strongest predictor of perceived health for female faces, with facial adiposity failing to reach statistical significance.

As mentioned previously, for a specific facial cue to serve as a valid cue to health, people need to reliably detect the cue and agree on its relationship to health or attractiveness. Some studies have reported that participants are able to reliably estimate a person's body fat percentage or body mass index (BMI) from facial cues alone (Fisher et al., 2013; Tinlin et al., 2013; Han et al., 2016; Phalane et al., 2017). The fact that participants are able to accurately detect changes in BMI likely stem from the fact that changes in BMI have been linked to a predictable set of morphological characteristics such as width-to-height ratio, perimeter-to-area ratio, cheek-to-jaw-width ratio (Coetzee et al., 2010), shape of lower face outline, nose width, and eyebrow position (Mayer et al., 2017). A study by Re et al. (2013) reported that people can detect changes in BMI as small as 1.3 kg/m² in male faces and 1.6 kg/m² in female faces from facial cues alone. Another study by Re and Rule (2016b) corroborated this finding by reporting that an average change in BMI of 1.33 kg/m² was sufficient for participants to report a noticeable change in the appearance of faces. In recent years there has also been a concerted effort from researchers to develop computer vision methods (Wen and Guo, 2013; Kocabey et al., 2017; Barr et al., 2018) and statistical models (Wolffhechel et al., 2015; Stephen et al., 2017) to predict BMI from facial images.

Being able to draw inferences about a person's weight from facial cues of adiposity, provides a potentially robust perceptual link between facial cues and actual health. Body fat percentage and BMI has consistently been linked to various negative health outcomes. Overweight individuals are subject to unfavorable health outcomes including diabetes, mental health problems, impaired immune function, high and low blood pressure and heart disease (Kopelman, 2007; Dixon, 2010; Zaccardi et al., 2017); while being severely underweight could be an indication that a person is vulnerable to communicable diseases (Fisher et al., 2014b), since being underweight is linked to malnutrition

and compromised immune function (Ritz and Gardner, 2006; Dobner and Kaser, 2018). Accurate estimation of overall body weight from facial cues can thus serve as an effective cue to health, as lower or higher than average body weight could be indicative of a person's past, current and future health problems (Coetzee et al., 2009; Henderson et al., 2016). In addition to the strong relationship between BMI and health outcomes, a study by Levine et al. (1998) has also established a strong link between cheek adipose tissue and visceral abdominal fat and a study by Lee and Kim (2014) established a correlation between the distances between the inferior ear lobes and visceral fat. Inferring visceral fat from facial cues is a potentially important predictor of health outcomes, as visceral fat has been connected to negative health outcomes due to its high metabolic activity (Després and Lemieux, 2006). BMI is also highly heritable, as twin studies have estimated the heritability of BMI to be as high as 0.57 in people between 20 and 29 years of age, even after environmental variation and cultural-geographic regions were accounted for (Silventoinen et al., 2017).

Tables 1, 2 present summaries of studies investigating the relationship between facial adiposity, attractiveness, and perceived health.

While research into the link between facial adiposity and health and attractiveness has been gaining momentum over the last decade, there are some challenges faced by researchers. Investigating the relationship between facial adiposity, actual health, and attractiveness provides a series of methodological challenges. As both underweight and overweight individuals are likely to be judged less healthy and attractive, Coetzee et al. (2009) proposed that the relationship between perceived health, attractiveness and adiposity is best represented by a quadratic relationship. Seeing as both overweight and significantly underweight individuals are subject to negative health outcomes, any large deviations from average levels of adiposity is likely to lead to decreases in health and attractiveness ratings. Subsequent studies confirmed that curvilinear models are often the best fitting models for these data (Phalane et al., 2017; Windhager et al., 2018), though this is not always the case (Coetzee et al., 2012; see for example Foo et al., 2017). In studies where sampling range restrictions occur for BMI, it is also possible that the range restriction obscures the quadratic pattern of the results. One notable study by Windhager et al. (2013) created a set of female adolescent geometric morphs that allowed the researchers to produce a set of 5 images which ranged from -5 SD (19% body fat) to $+5$ SD (64% body fat). These facial morphs were then used in a recent study where 274 male and female participants had to rate each of the 5 images on maturity, dominance, masculinity, perceived health and attractiveness (Windhager et al., 2018). Analysis revealed that both health and attractiveness consistently produced distinct curvilinear shapes across all rater categories and genders.

In addition to the complex statistical relationship that exists between facial adiposity and preference ratings, there is also evidence to suggest that humans integrate information from multiple facial cues to evaluate faces. One of the facial cues that

TABLE 1 | Summary of studies that investigated the relationship between facial adiposity and attractiveness.

Study (First Author)/Year	Stimuli			Results
	N	Gender	Mean Age	
Coetzee et al., 2009	84	M + F	<30	Significant curvilinear relationship between facial adiposity and attractiveness ($R^2 = 0.186$)
Klaczynski et al., 2009	64	M + F	<18	Average weight facial drawings were considered more attractive than obese facial drawings ($\eta_p^2 = 0.91$)
Coetzee et al., 2011	96	M + F	<30	On average, female faces were transformed by female raters to represent a BMI of 19.76 kg/m ² to reflect optimal attractiveness On average, female faces were transformed by male raters to represent a BMI of 20.01 kg/m ² to reflect optimal attractiveness
Coetzee et al., 2012	45	F	<30	Significant negative relationship between adiposity and attractiveness ($r = -0.216$)
Rantala et al., 2013a	52	F	<20	Negative correlation between facial adiposity and attractiveness ($r = -0.481$) No significant relationship between Hepatitis B antibody response and attractiveness ratings
Fisher et al., 2013*	100	M + F	<30	Small negative correlation between facial adiposity and attractiveness for males ($r = -0.14$) and females ($r = -0.19$)
Re and Perrett, 2014	20	M + F	<30	On average, female faces were transformed to represent a BMI of 18.19 kg/m ² to reflect optimal attractiveness On average, male faces were transformed to represent a BMI of 22.46 kg/m ² to reflect optimal attractiveness Female faces were also transformed to a significantly lower BMI to maximize attractiveness compared to male faces ($\eta_p^2 = 0.78$)
Batres and Perrett, 2014	10	M + F	<30	Participants without internet access preferred faces with higher adiposity for females ($\eta_p^2 = 0.170$), but not for males ($\eta_p^2 = 0.0005$)
Fisher et al., 2014b	160	M + F	-	Two way interaction revealed that facial adiposity had a larger impact on facial attractiveness ratings ($d = 4.7$) compared to health judgments ($d = 3.76$)
Fisher et al., 2014a*	100	M + F	<30	Small negative correlation between facial adiposity and attractiveness for males ($r = -0.16$) and females ($r = -0.13$)
Han et al., 2016	96	F	<30	Significant negative relationship between adiposity and attractiveness ($r = -0.59$)
Re and Rule, 2016b	40	M + F	<30	On average, female faces were transformed to represent a BMI of 19.11 kg/m ² to reflect optimal attractiveness On average, male faces were transformed to represent a BMI of 23.79 kg/m ² to reflect optimal attractiveness
Foo et al., 2017	101	M	<30	Significant negative relationship between facial adiposity and attractiveness ($r = -0.50$)
	80	F	<30	Significant negative relationship between facial adiposity and attractiveness ($r = -0.58$)
Batres et al., 2017	10	M + F	<30	Participants from rural areas in Malaysia preferred faces with higher adiposity for females ($d = 0.69$), but not for males ($d = 0.15$). Participants from rural areas in El Salvador preferred faces with higher adiposity for females ($d = 0.97$), but not for males ($d = 0.13$).
Phalane et al., 2017	92	M	<30	Significant curvilinear relationship between adiposity and perceived health ($R^2 = 0.138$)
Windhager et al., 2018	-	F	-	Curvilinear relationship between facial adiposity and attractiveness for male raters: adolescents ($R^2 = 0.22$), young adults ($R^2 = 0.44$) and older adults ($R^2 = 0.29$).
	-	F	-	Curvilinear relationship between facial adiposity and attractiveness for female raters: adolescents ($R^2 = 0.24$), young adults ($R^2 = 0.46$) and older adults ($R^2 = 0.33$).

*Used the same facial stimuli panels

have been identified as playing a particularly important role in mediating the relationship between facial adiposity and health ratings is skin condition. According to Fisher et al. (2014b) judgments of health and attractiveness from facial adiposity in isolation produces a conundrum for raters, since it becomes very difficult to differentiate between individuals with low levels of facial adiposity due to health and those that have low facial adiposity levels due to ill-health. One way around this problem is to integrate information from skin condition with facial adiposity to provide more information on the potential risk

associated with the facial adiposity levels being perceived. An experiment conducted by Fisher et al. (2014b) provided evidence that participants likely integrate both skin color and facial adiposity cues when judging faces for health and attractiveness. Redness and yellowness in the face has been associated with perceptions of health (Stephen et al., 2009, 2011) and in the study by Fisher et al. (2014b) there was evidence to suggest that participants relied more on color cues to decide if a face looked healthy or attractive when facial adiposity levels were relatively low.

TABLE 2 | Summary of studies that investigated the relationship between facial adiposity and perceived health.

Study (First Author)/Year	Stimuli			Results
	N	Gender	Mean Age	
Coetzee et al., 2009	84	M + F	<30	Significant curvilinear relationship between adiposity and perceived health ($R^2 = 0.263$)
Coetzee et al., 2011	96	F	<30	Female raters indicated significantly lower BMI is required for attractiveness (19.76 kg/m ²) compared to optimum health (20.84 kg/m ²). No such difference was found for males
Fisher et al., 2014b	20	M + F	-	Facial adiposity had a significant effect on perceived health ratings ($d = 3.76$)
Han et al., 2016	96	F	<30	Significant linear relationship between adiposity and perceived health ($r = -0.52$)
Stephen et al., 2017	60	M + F	-	Raters significantly reduced BMI ($d = 3.56$) and percentage body fat ($d = 3.96$) to enhance healthy appearance
Axelsson et al., 2018	16	M + F	-	Higher sickness ratings were related to a more swollen facial appearance ($Z = 4.1$, $p < 0.001$)
Henderson et al., 2016	118	M + F	<30	Significant quadratic relationship between adiposity scores and perceived health ratings ($R^2 = 0.25$)
	67	F	<30	Significant quadratic relationship between adiposity scores and perceived health ratings ($R^2 = -0.35$)
Foo et al., 2017	101	M	<30	Significant negative relationship between facial adiposity and perceived health ($r = -0.56$)
	80	F	<30	Significant negative relationship between facial adiposity and perceived health ($r = -0.40$)
Phalane et al., 2017	92	M	<30	Significant curvilinear relationship between adiposity and perceived health ($R^2 = 0.138$)
Windhager et al., 2018	5	F	-	Curvilinear relationship between facial adiposity and perceived health for male raters: adolescents ($R^2 = 0.13$), young adults ($R^2 = 0.28$) and older adults ($R^2 = 0.13$)
	5	F	-	Curvilinear relationship between facial adiposity and perceived health for female raters: adolescents ($R^2 = 0.13$), young adults ($R^2 = 0.30$) and older adults ($R^2 = 0.18$)

According to the World Health Organization (WHO) classification, a healthy BMI for an average adult is between 18.5 and 24.9 kg/m² (World Health Organization, 2018). Numerous studies have shown that people generally prefer facial adiposity levels that are associated with the normal range of the World Health Organization's classification for BMI and that lower levels of facial adiposity is typically rated as more attractive, even across cultures (Coetzee et al., 2011, 2012, 2014; Re et al., 2011; Re and Rule, 2016b). For example, a study by Coetzee et al. (2011) found that, female participants judging other female faces, indicated that facial adiposity levels associated with a BMI of 19.76 kg/m² are optimally attractive, while male raters indicated that adiposity levels associated with a BMI of 20.01 kg/m² are optimally attractive for female faces. A similar study by Re and Rule (2016b) found that participants preferred facial adiposity levels associated with a BMI of 19.11 kg/m² for female faces and 23.79 kg/m² for male faces. Coetzee (2011) compared facial adiposity preferences between a cohort of British and African participants and found that Caucasian males, African females and African males all preferred similar levels of facial adiposity for attractiveness and health. Despite the evidence that there is agreement across cultures on facial adiposity levels associated with health and attractiveness, there does appear to be some degree of divergence in BMI preference judged from facial adiposity and BMI preferences as judged from bodies across cultures. For example, a study by Tovée et al. (2006) found that ethnic South African Zulus preferred a BMI that is above the upper limit of the normal range for both health and attractiveness when judging female bodies. Since most studies on

facial adiposity preferences utilize university students, Coetzee et al. (2012) argues that this discrepancy could potentially be attributed to repeated exposure to Western media ideals and improved access to resources that African university students enjoy. This line of reasoning is also supported by Tovée et al. (2006) who found that South African born Zulus who immigrated to the United Kingdom preferred a lower BMI for both attractiveness and health compared to Zulus still residing in South Africa (also see Tovée et al., 2007). At present, more research is needed to elucidate the cultural and environmental factors associated with facial adiposity preferences across cultures.

There also appears to be relative differences between males and females when it comes to judgments of facial adiposity levels required for a healthy appearance for female faces. Coetzee et al. (2011) report that females preferred a slightly lower BMI (19.76 kg/m²) for attractiveness compared to optimum health (20.84 kg/m²) when rating other females, while no such difference was found for males. One hypothesis that has been put forth to account for this discrepancy, is that socio-cultural factors related to judgments of body fat influences and shapes female body weight ideals, especially via media exposure to a thin ideal (Groesz et al., 2002; Harrison and Hefner, 2006; Coetzee et al., 2011). Though the link between facial adiposity and perceived health and attractiveness appears to be remarkably stable, facial adiposity likely interacts in complex ways with various biological, environmental and socio-cultural factors, as well as other facial cues such as symmetry, skin color, skin texture, sexual dimorphism and averageness in shaping social and health perceptions derived from faces.

Facial Adiposity and Health Outcomes

One of the primary reasons why researchers are interested in facial adiposity is its potential to act as a robust cue to health outcomes. Given the strong relationship that exists between facial adiposity and BMI, negative health outcomes related to BMI should therefore also be linked to facial adiposity. To date, numerous studies have found a relationship between facial adiposity and actual health outcomes, although the results are not always consistent.

Cardiovascular Health

A study by Coetzee et al. (2009) found that facial adiposity was associated with systolic blood pressure ($r = 0.28$) and diastolic blood pressure ($r = 0.43$), as well as a cardiovascular-illness component ($r = 0.41$). In addition to the results reported by Coetzee et al. (2009) a study by Reither et al. (2009) also found a link between perceived facial adiposity and cardiovascular health. Reither et al. (2009) used 3,027 yearbook photographs of teenagers from the Wisconsin Longitudinal Study (WLS), which also captured basic health data from 1957 to 2004. The researchers found that increased facial adiposity ratings in adolescent faces were associated with high blood pressure ($OR = 1.25$), heart trouble ($OR = 1.20$) and heart disease ($OR = 1.69$) later in life. Recently a study by Stephen et al. (2017) also reported that a geometric morphometric model that captured aspects of facial shape from images, could be used to predict 21% of the variance in blood pressure in their sample. These results indicate that there is a robust link between cardiovascular health and facial adiposity.

Immunity

Coetzee et al. (2009) reported that higher levels of rated facial adiposity was associated with an increase in the number of colds and flu's participants reported ($r = 0.209$), bout length of those reported colds and flu's ($r = 0.24$), the frequency of antibiotics use reported by participants ($\chi^2 = 11.706$) and a respiratory-illness component ($r = 0.29$). These relationships remained significant, even after the researchers controlled for parental income and age. A study on male participants by Rantala et al. (2013a) found an association between perceived facial adiposity and immunity ($\beta = -0.38$) using a direct measure of immune function (antibody response to a Hepatitis B vaccination). The researchers also reported that perceived facial adiposity mediated the relationship between antibody response and facial attractiveness. However, a similar study using female participants, found no link between the antibody response produced by a Hepatitis B vaccination and percentage body fat, though it should be mentioned that no direct measures of facial adiposity were included in the study (Rantala et al., 2013b). Although the studies reported above found indications that immune function may be linked to facial adiposity, especially in males, other studies have not produced the same pattern of results when using different measures of immune function. For example, a study by Phalane et al. (2017) using African participants found no significant relationship between

adiposity and immune responsiveness as measured by functional cytokine profile and C-reactive proteins in African men. It is worth mentioning, however, that the sample size for the cytokine profile was relatively low ($n = 41$) and could thus reflect a potential type II error due to a lack of statistical power. A recent study by Foo et al. (2017) also found no relationship between facial adiposity and immune function (bacterial killing capacity, overall bacterial immunity, bacterial suppression capacity, and lysozyme activity) in either males or females. To date, there is some evidence to suggest that facial adiposity and immune function are linked, especially in men, although more research is needed to clarify the potential relationship between different aspects of immune function and facial adiposity.

Mental Health

Relatively few studies have investigated the link between facial adiposity and mental health. A study by Martinson and Vasunilashorn (2016) that was conducted using the same basic methodology applied to the WLS data used by Reither et al. (2009), found that females who were rated as being overweight ($M = 8.79$) in 1957 scored significantly higher on the Center for Epidemiological Studies Depression Scale (CES-D) later in life compared to normal weight females ($M = 6.94$). No statistically significant difference was found between overweight ($M = 6.40$) and normal weight ($M = 5.97$) CES-D scores for male students. Additionally, a study conducted by Tinlin et al. (2013) using female participants, also found that rated facial adiposity was negatively correlated with a psychological condition factor ($r = 0.29$). These findings suggest that facial adiposity may also serve as a valid cue to mental health in females, although more research is required to draw any conclusions about the link between mental health and facial adiposity in males.

Hormones

The relationship between facial adiposity and various hormone levels has also produced largely mixed results. For example, a study by Tinlin et al. (2013) conducted using female participants reported a significant negative correlation between salivary progesterone levels ($r_s = -0.30$) and facial adiposity, but found no relationship between facial adiposity and estradiol levels, though the results should be interpreted with caution due to the small sample size ($n = 49$). Rantala et al. (2013a) found a significant relationship between circulating testosterone and facial adiposity in men ($r = -0.52$). Han et al. (2016) found no relationship between facial adiposity and cortisol levels in females, while Rantala et al., 2013b also found no relationship between cortisol levels and percentage body fat in females. To date, most of the research done on the relationship between hormones and facial adiposity has been done using female participants.

Other Health Outcomes

In addition to cardiovascular health, immune function, mental health and hormone levels, some studies also investigated potential links between facial adiposity and other health outcomes such as diabetes and arthritis or indicators of health

such as oxidative stress and sperm health. For example, Reither et al. (2009) found that adolescents who were rated as having higher levels of facial adiposity were also more likely to experience muscle aches ($OR = 1.13$), shortness of breath ($OR = 1.10$) and chest pains ($OR = 1.21$) during the course of their lives, but also showed an increased risk of developing arthritis ($OR = 1.19$), diabetes ($OR = 1.44$), and die of non-accidental causes ($OR = 1.32$). A study by Foo et al. (2017) investigated the relationship between facial adiposity and oxidative DNA damage, as well as level of lipid peroxidation. The researchers found a significant relationship ($r = -0.22$) between facial adiposity and urinary 8-hydroxy-2'-deoxyguanosine (8-OHdG), a biomarker for oxidative stress and carcinogenesis (Valavanidis et al., 2009) for males, but not for females. However, Foo et al. (2017) also reported that no relationship existed between facial adiposity and a measure of lipid peroxidation (isoprostane levels) in either males or females. Lastly, Foo et al. (2017) found no relationship between facial adiposity and sperm health (rapid progressive motility, linearity of sperm movement, sperm concentration and percentage motile sperm) for males.

A study by Tinlin et al. (2013) investigated the relationship between facial adiposity and general health markers in female participants. The researchers found that rated facial adiposity was not correlated with a physical condition factor, although the researcher did note that there was evidence for high collinearity between the physical condition and psychological condition factors in their study. A combination of these factors (labeled general condition factor), displayed a moderate association with facial adiposity ($r = 0.41$).

When combined, the results from these studies indicate that facial adiposity can serve as a reliable indicator of actual health outcomes including immune function, cardiovascular health, respiratory health, mental health and oxidative stress, although more research is needed to elucidate the link between these health outcomes and facial adiposity. The next section briefly touches upon factors that can potentially influence or mediate the relationship between facial adiposity and judgments of attractiveness and health.

Table 3 presents a summary of studies investigating the relationship between facial adiposity and health outcomes.

TABLE 3 | Summary of studies that investigated the relationship between facial adiposity and health outcomes.

Study (First Author)/Year	Stimuli			Result
	N	Gender	Mean Age	
Coetzee et al., 2009	84	M + F	<30	Significant relationship between facial adiposity and cold and flu Number ($r = 0.21$), bout length ($r = 0.24$), frequent antibiotics use ($\chi^2 = 11.71$), respiratory-illness component ($r = 0.29$), systolic blood pressure ($r = 0.28$), diastolic blood pressure ($r = 0.43$), cardiovascular-illness component ($r = 0.41$)
Reither et al., 2009	3027	M + F	<30	Adolescents who were rated as having higher levels of facial adiposity were more likely to experience muscle aches ($OR = 1.13$ [95% CI = 1.02, 1.26]), shortness of breath ($OR = 1.10$ [95% CI = 0.93, 1.31]) and chest pains ($OR = 1.21$ [95% CI = 0.96, 1.53]). Adolescents who were rated as having higher levels of facial adiposity were also more likely to develop arthritis ($OR = 1.19$ [95% CI = 1.08, 1.31]), high blood pressure ($OR = 1.25$ [95% CI = 1.11, 1.41]), diabetes ($OR = 1.44$ [95% CI = 1.09, 1.90]) and heart trouble ($OR = 1.20$ [95% CI = 0.96, 1.51]) and die of non-accidental causes ($OR = 1.32$ [95% CI = 1.10, 2.58]) or heart disease ($OR = 1.69$ [95% CI = 1.10, 2.58]).
Tinlin et al., 2013	50	F	<30	Significant correlation between facial adiposity and a psychological condition factor ($r = 0.29$), as well as salivary progesterone ($r_s = -0.30$). No relationship between facial adiposity and a physical condition factor ($r = 0.26$) or estradiol levels ($r_s = -0.18$)
Rantala et al., 2013a	69	M	<30	Circulating testosterone was positively correlated with adiposity ($r = 0.52$) Hepatitis B anti-body response was also significantly associated with facial adiposity ($\beta = -0.38$)
Han et al., 2016	96	F	<30	No Significant relationship between adiposity and average cortisol levels ($r = 0.05$)
Martinson and Vasunilashorn, 2016	4410	M + F	<30	Females who were rated as being overweight ($M = 8.79$) scored significantly higher on the CES-D depression inventory compared to females who were rated as normal weight ($M = 6.94$) There was no statistically significant difference between CES-D scores for overweight ($M = 6.40$) and normal weight ($M = 5.97$) males
Foo et al., 2017	101	M	<30	Significant negative correlation between facial adiposity and an indicator of oxidative stress (8OHdG levels) ($r = -0.22$). No relationship between facial adiposity and an indicator of lipid peroxidation, immune function (bacterial killing capacity, overall bacterial immunity, bacterial suppression capacity and lysozyme activity), or semen quality (rapid progressive motility, linearity of sperm movement, sperm concentration and percentage motile sperm)
	80	F	<30	No relationship between facial adiposity and indicators of oxidative stress and lipid peroxidation or immune function (bacterial killing capacity, overall bacterial immunity, bacterial suppression capacity and lysozyme activity)
Phalane et al., 2017	92	M	<30	No significant relationship between adiposity and immune function as measured by a cytokine component ($r = 0.282$) or C-reactive protein ($r = -0.0007$).

Factors That Potentially Contribute to Adiposity Preferences

While numerous studies have found a robust link between facial adiposity and perceptions of health and attractiveness, a number of potential mediating factors have been identified that could serve to modify our facial adiposity preferences. A study by Re et al. (2011) demonstrated that our preferences for particular adiposity levels can be altered, at least in the short term. First, the researchers used a pre-test to establish a baseline for participants' facial adiposity preferences. By exposing participants to either "plus-sized" female bodies or regular female bodies, the researchers could induce an after-effect where participants who viewed the "plus-size" bodies increased their adiposity preference by an average of 0.5 kg/m². Participants thus preferred faces higher in adiposity after viewing the "plus-size" bodies. This finding illustrates an adaptation effect, whereby exposure to certain stimuli can create a preference shift toward that particular type of stimulus. According to the authors, this after-effect could be an indication that overall attractiveness judgments require integration of multiple cues, though these cues are often investigated independently.

The idea of an adaptation effect is also related to repeated exposure to facial cues found in our own ethnicity, or ethnicities that we are regularly exposed to. In a study by Schneider et al. (2013) Japanese and German participants were asked to provide body weight estimates for both German and Japanese individuals based on facial photographs alone. The researchers found that German observers tended to slightly overestimate the body weight of the Japanese faces, while the Japanese participants significantly underestimated the body weight of the German faces. The researchers argue that one potential explanation for the gross underestimation of body weight for German faces by Japanese participants, is that Japanese participants utilized an inappropriate reference point regarding the link between facial shape and body weight for German faces. Since Japanese faces tend to display more rounded and broad (brachycephalic) head proportions, German faces associated with an average body weight appeared more slender, compared to a Japanese face associated with an average body weight, which Japanese participants would have been more familiar with. These findings imply that accurate weight perceptions derived from facial cues, is also contingent upon reasonable exposure and familiarity with the association between facial phenotypic variation and body weight of a particular ethnic group.

Another factor that can play a significant role in influencing people's adiposity preferences, is potential pathogen exposure and how it relates to our behavioral immune system. Studies have shown that levels of pathogen disgust sensitivity can have a dramatic influence on individuals' perceptions of facial cues associated with attractiveness or perceived health (Park et al., 2012; Jones et al., 2013). These results suggest that a pathogen disgust reaction may indeed play a role in our relative tolerance for facial cues that may signal ill-health. A study by Fisher et al. (2013) found that men with higher pathogen disgust responses showed preferred facial cues associated with lower weight, indicating that individual differences in pathogen disgust

also play a role in relative levels of facial adiposity that people find attractive or perceive to be healthy. These results are also supported by Rantala et al. (2013a) who found that facial adiposity mediated the association between facial attractiveness and immune response in men. However, a study by Dixon et al. (2017) found no relationship between facial adiposity preferences and malarial prevalence for people living in urban or rural areas in Vanuatu, a series of small Pacific islands. These results suggest that the relationship between facial adiposity preferences and pathogen exposure or disgust, while theoretically elegant, have to be interpreted with caution.

Environmental pressures also appear to play a role in shaping our preferences for facial or body adiposity. For example, it has been demonstrated that resource scarcity could play a role in males' preference of breast size, with men from lower socio-economic regions preferring larger breasts, due to breast size being a good indicator of adipose tissue reserves (Dixon et al., 2011; Swami and Tovée, 2013). A study by Batres and Perrett (2014) found that people in rural areas of El Salvador tend to prefer female faces with higher levels of adiposity, compared to people from urban areas. One hypothesis that has been proposed to account for this difference is the fact that rural areas, especially in developing nations, tend to be poorer and people from these areas tend to have less access to resources such as food or medicine. A similar study conducted by Swami and Tovée (2007) also provide evidence that people from resource-poor areas of Thailand preferred female bodies that were higher in BMI compared to people from more industrialized areas of Thailand. A more recent study by Batres et al. (2017) replicated the results of Batres and Perrett (2014) in another cohort of people from El Salvador, as well as Malaysia. What is especially interesting about these results is the fact that a difference in preference for facial adiposity levels between more developed and poorer areas was only found for females' faces. One possible reason for this is that a trade-off exists in poorer areas, where males have to weigh the potential negative long-term health outcomes associated with higher levels of adiposity and more immediate concerns regarding survival and reproductive fitness of females (Batres et al., 2017). Another possible contributor to this phenomenon could be increased media exposure to thin ideals in more developed areas, which then serves to alter people's baseline expectations of attractiveness or health, via an adaptation effect as described by Re et al. (2011) and Batres and Perrett (2014). It should be pointed out, however, that there are studies that problematize the view that people in resource-poor areas prefer bodies or faces associated with a slightly higher BMI. For example, a study by Dixon et al. (2017) found no evidence to support the hypothesis that males from rural areas would prefer females' faces associated with higher levels of facial adiposity. A series of studies conducted in China, Papua New Guinea, Cameroon, Indonesia, Samoa and New Zealand, also found a high degree of cross-cultural consensus when participants were asked to judge female body attractiveness, with both males and females preferring a lower waist-to-hip ratio (WHR) regardless of BMI (Dixon et al., 2010a,b; Singh et al., 2010). These results suggest that sexually dimorphic fat distribution could in fact be

more important than overall BMI when judging female body attractiveness, even across ethnicities or socio-economic regions.

A study by Weston et al. (2015) revealed that judgments of weight in faces can even be influenced by the expression of the facial stimuli. The researchers observed that participants tended to rate male faces containing sad expressions as more overweight compared male faces containing a neutral facial expression. A study by Henderson et al. (2016) also found that downward mouth curvature was negatively correlated with perceived health ($r = -0.20$) and upward mouth curvature was positively correlated with perceived health ($r = 0.51$), presumably because a frown is associated with a sad facial expression that reflect a negative mood or mental state. It is thus likely that facial cues are also evaluated within the context of emotional attributions, adding yet another layer of complexity to the way in which we judge physical and emotional health from facial cues in others.

Methodological Considerations in the Study of Facial Adiposity

In the laboratory there are several factors related to the facial stimuli that can influence the outcome of a study. For example, a study by Jones et al. (2012) found that perceptions and the derived judgments of facial stimuli can be altered by changing the viewing angle of the face. According to the authors, this effect is likely attributable to directional asymmetries in the perception of facial cues related to shape. A study by Russell et al. (2016) found that facial contrast can also influence perceptions of health within the face, with faces low in contrast being judged as less healthier. These two studies highlight the importance of controlling for subtle factors such as viewing angle and facial contrast in the generation of facial stimuli, as they could potentially confound the results of the study.

In addition to the way in which stimuli are generated, statistical and research design considerations also need to be taken into account when conducting facial morphology research. For example, Windhager et al. (2018) proposes that a major drawback of the traditional reliance on p -values in the analysis of data derived from facial stimuli, is the fact that facial morphology research often involves complex nested designs. Most designs in the field use human raters to rate multiple facial stimuli on one or more dimensions. Researchers then often calculate a mean rating for a particular face and although inter-rater reliability is mostly high ($\alpha > 0.80$), the variance associated with rater responses is usually lost when rater responses are collapsed to a single mean. Many studies also use suboptimal sample sizes ($n > 100$) due to the complexity of the designs, which can lead to a lack of statistical power in detecting smaller effect sizes (see for example Tinlin et al., 2013; Phalane et al., 2017). In reality, collinearity concerns are likely always going to be a potential problem researchers will have to deal with, due to the complex facial cue integration that happens within an organic judgment faces. However, recent advances in the creation of facial morphs can potentially allow researchers to manipulate facial cue dimensions with increasing precision, while also theoretically isolating specific facial cues within a panel of facial stimuli

(Windhager et al., 2013, 2018; Re and Perrett, 2014). Bottom-up or data-driven methodologies can also offer researchers a way to increase the precision with which facial cues are specified by reducing shape coordinates to a set of principle components that can be used in further analysis (Wolffhechel et al., 2015; Henderson et al., 2016).

Estimating BMI From Facial Adiposity

One of the key components of utilizing weight related facial cues to make inferences about people's health is that people should be able to accurately judge body mass from facial cues alone. To our knowledge no study has provided a quantitative synthesis of the link between facial adiposity and judgments of body weight across different studies published in the field. A meta-analysis was therefore conducted to evaluate the strength of the relationship between judgments of facial adiposity and BMI/ percentage body fat.

METHODS

Inclusion Criteria

All peer-reviewed journal articles that provided an effect size that quantifies the relationship between perceived facial adiposity and estimations of BMI or percentage body fat were included in the meta-analysis. Studies that reported more than one effect size based upon the same facial stimuli were excluded from the analysis due to potential biases in heterogeneity estimates. For example, a study by Coetzee et al. (2012) calculated a correlation coefficient for the relationship between perceived facial adiposity and BMI, as well as percentage body fat for the same participants. Although two effect sizes can be extracted from this study, the two effect sizes obtained from the same group would not be statistically independent. Similarly, Fisher et al. (2014a) used the same facial stimuli as Fisher et al. (2013), thus the effect sizes obtained from these two studies would also not be statistically independent. The effect size for the relationship between facial adiposity and percentage body fat from Coetzee et al. (2012) and the study by Fisher et al. (2014a) were thus excluded from the analysis. No restriction criteria were placed on the age, gender or ethnicity of either the facial stimuli or raters in studies to be included in the meta-analysis.

Search Strategy

A systematic review protocol was developed in accordance with the Preferred Reporting for Items for Systematic Reviews and MetaAnalyses-Protocols (PRISMA) (Moher et al., 2009). The PRISMA protocols can easily be reproduced by other researchers thereby ensuring transparency and reliability of meta-analyses (Gurevitch et al., 2018). The use of PRISMA protocols is becoming increasingly popular within the field of evolutionary psychology (Geniole et al., 2015; see for example Gouda-Vossos et al., 2018). In line with the protocol developed for this study, all studies that reported an effect size for the relationship between perceived facial adiposity and BMI or percentage body fat were identified by searching relevant databases for peer-reviewed articles that contained the phrase "facial adiposity": Pubmed (9), EbscoHost (27), Science Direct (66), ProQuest

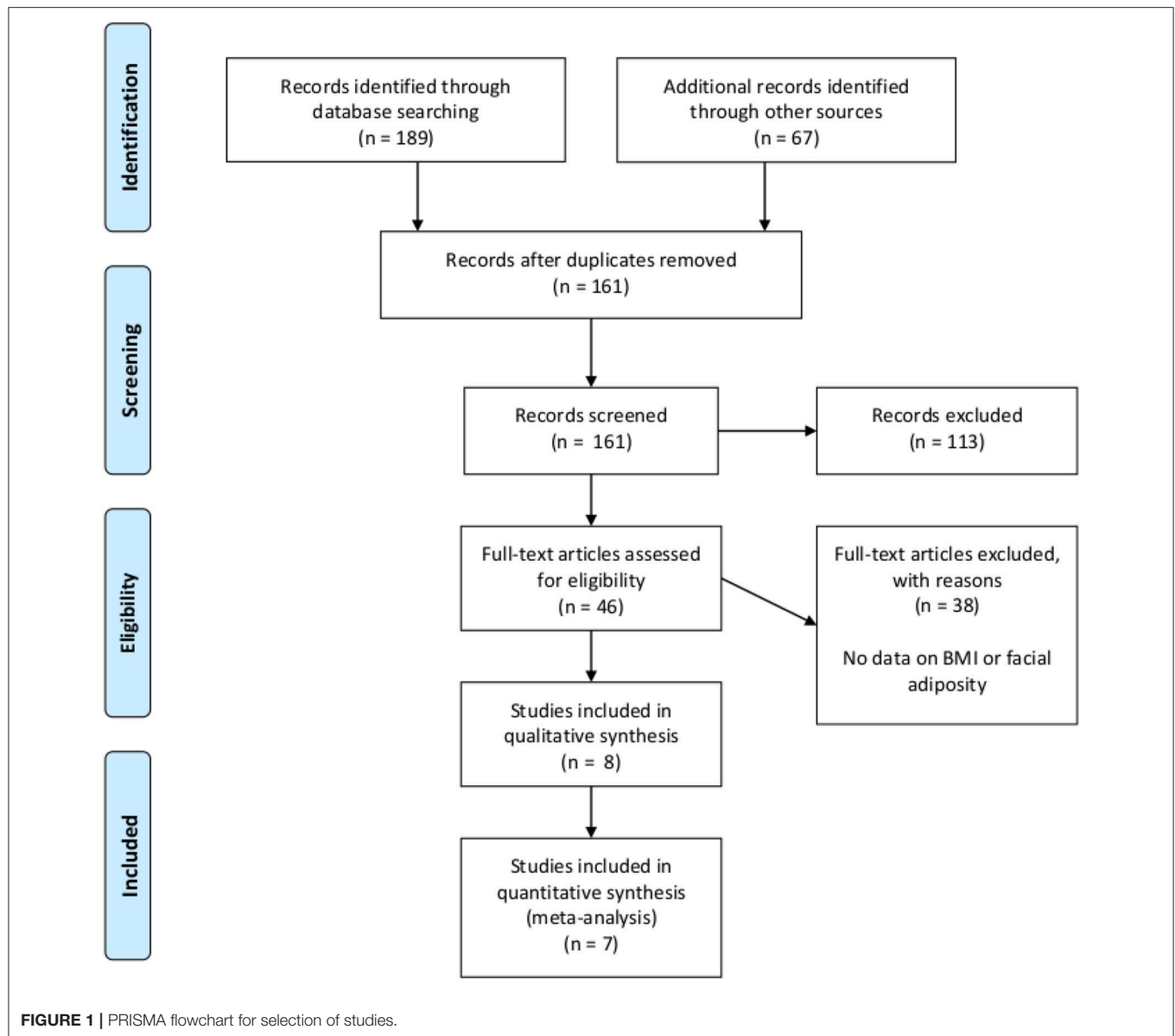


TABLE 4 | Summary of studies that investigated the relationship between facial adiposity and BMI/Percentage body fat.

Study (First Author)/Year	Stimuli				Raters			Effect Size (<i>r</i>)
	<i>n</i>	Gender	Age	BMI	<i>n</i>	Gender	Age	
Coetzee et al., 2009	84	M + F	<i>M</i> = 21.13	<i>M</i> = 22.95	55	M + F	-	0.66
Coetzee et al., 2012	45	F	<i>M</i> = 19.84	-	30	M + F	-	0.75
Tinlin et al., 2013	50	F	<i>M</i> = 20.3	<i>M</i> = 23.2	21	M + F	<i>M</i> = 21.1	0.68
	50	F	<i>M</i> = 24.3	<i>M</i> = 20.1	26	M + F	<i>M</i> = 23.2	0.66
	50	F	<i>M</i> = 19.74	<i>M</i> = 23.8	160	M + F	<i>M</i> = 23.77	0.63
Rantala et al., 2013b	69	M	<i>M</i> = 23	<i>M</i> = 17.36	14	F	<i>M</i> = 23.6	0.75
Fisher et al., 2013	50	M	<i>M</i> = 24.2	<i>M</i> = 20.1	50	M + F	<i>M</i> = 22.54	0.58
	50	F	<i>M</i> = 24.3	<i>M</i> = 20.1	50	M + F	<i>M</i> = 24.11	0.66
Han et al., 2016	96	F	<i>M</i> = 21.42	<i>M</i> = 25.53	463	M + F	<i>M</i> = 24.14	0.67
Phalane et al., 2017	92	M	<i>M</i> = 20.4	<i>M</i> = 21.7	20	F	<i>M</i> = 22.5	0.68

(66), and Web of Science (21). Once databases were searched, reference lists of prominent published articles were also searched and any additional relevant abstracts were added. Results were entered into the open-source software management package Zotero Version 5.0. Duplicate entries were removed, after which abstracts were screened for relevance in line with the inclusion criteria by the authors. After the initial screening procedure, the full-text articles were downloaded for the remaining list and re-screened for relevance. See **Figure 1** for the PRISMA flow diagram.

Data Extraction

A data extraction template was used to capture all the relevant data from the articles selected for inclusion in the meta-analysis. The data extraction sheets included detailed information on the types of stimuli including gender, age, BMI/percentage body fat, ethnicity, as well as the rating scale used and if the facial stimuli were captured in 2D or 3D. Demographic information on the raters was included as well, including gender, age and ethnicity. Finally, effect sizes were captured for all relevant studies.

Data Synthesis

All extracted effect sizes were converted to Pearson's r before any analyses were conducted. As discussed previously, studies often use the same facial stimuli or raters multiple times during a study, thus creating a complex nested design structure. This nested structure means that it is not unusual for a particular study to produce two or more analysis outcomes that were derived from highly correlated or identical sources within the study. If this statistical dependency is not adequately accounted for in the analysis procedures used by the researchers, they risk introducing bias in the variance estimation during the analysis (Hunter

and Schmidt, 2004). In order to account for the within-study statistical dependence, we aggregated all effect sizes within each study to produce a single estimated effect size for that study using the "Agg" function contained in the "Mac" package for R (Del Re and Hoyt, 2012; R Core Team, 2013). As information on the true correlations of within-study effect sizes were unavailable, all within-study effect size correlations were fixed at 0.50 (Wampold et al., 1997). This aggregation procedure produced seven effect size estimates, one corresponding to each study. Due to key differences between studies regarding selection and presentation of facial stimuli, as well as sample composition, meta-analyses were conducted using random-effect models (Borenstein et al., 2010; Quintana, 2015). The aggregated Pearson's r for each study was transformed to a Fisher's z scale, as r is not normally distributed. These effect size estimates were back transformed where appropriate in the presentation of the results. See **Table 4** for a summary of studies included in the meta-analysis.

RESULTS

A random effects model weighted by sample size was performed using the "metaphor" package in R (Viechtbauer, 2010). The analysis revealed a strong positive overall correlation between perceived facial adiposity ratings and BMI/percentage body fat [$r = 0.71$; 95% CI (0.66, 0.76), $p < 0.001$]. Heterogeneity statistics revealed no statistically significant between-study heterogeneity [$Q = 7.12$ (df = 6), $p = 0.31$; $I^2 = 19.36\%$; $\tau^2 = 0.004$; $\tau = 0.06$]. According to guidelines published by Higgins et al. (2003), an I^2 proportion below 25% is a good indicator of low between-studies variability and can serve as a more robust indicator of between-study variance in small sample sizes.

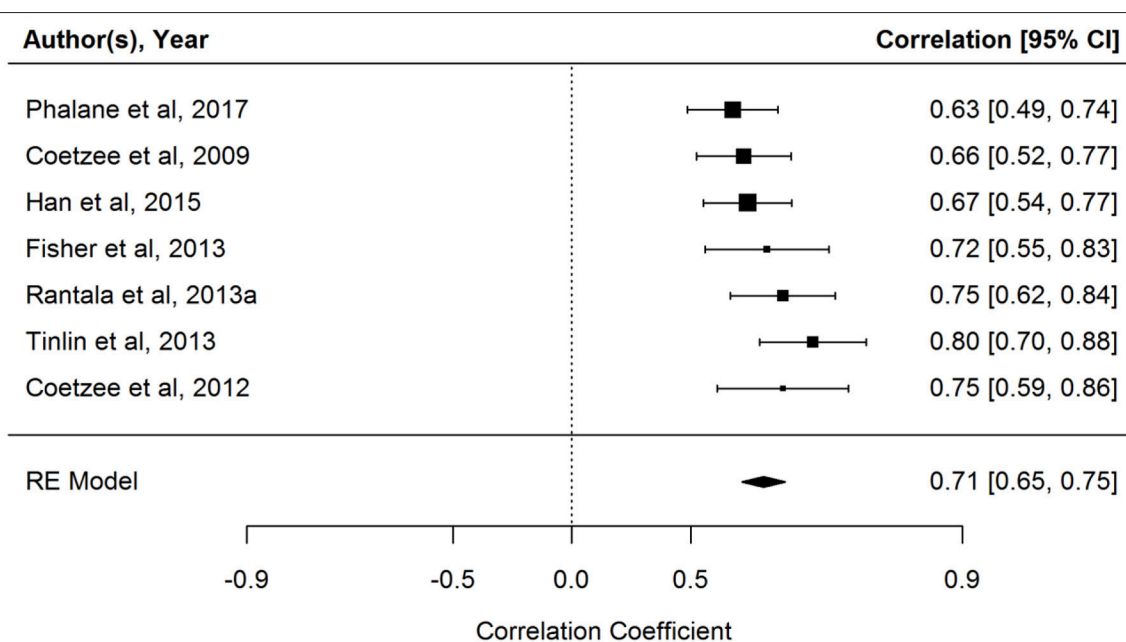


FIGURE 2 | Forest plot of relationship between perceived facial adiposity and BMI/ percentage body fat.

Sensitivity analysis was conducted to investigate if any of the studies included in the analysis contributed disproportionately to heterogeneity. Outlier and influential case diagnostics were conducted, which indicate that only Coetzee et al. (2012) (Study 6 in **Figure 2**) displayed signs of being a potential outlier or influential case (Viechtbauer and Cheung, 2010). See **Figure 3** for influential case diagnostics. To test if the removal of Coetzee et al. (2012) from the analysis would influence the overall model fit, the study was excluded and the model re-fitted. The re-fitted model did not make a significant impact on the interpretation of the overall results however [$r = 0.70$; 95% CI (0.65, 0.76), $p < 0.001$]. Moderator analysis also revealed that the ethnicity of the facial stimuli (African or Caucasian) did not have a moderating effect on the relationship between perceived facial adiposity ratings and BMI/ percentage body fat [$Q(1, 0.50, p = 0.48)$].

Risk of Bias

Figure 4 shows a funnel plot of correlation coefficients plotted against standard errors for each study. Egger's regression test showed no evidence of any asymmetry in the funnel plot. Due to the relatively low sample size ($n = 7$) of the study, visual symmetry was hard to estimate and therefore no trim-and-fill methods were used.

DISCUSSION

Facial adiposity has consistently been linked to perceptions of attractiveness and health, with heavier faces being judged to be more unattractive and unhealthier. To date, facial adiposity has also been linked to a number of actual health outcomes including: cold and flu number, duration of colds and flu, frequency of antibiotic use, respiratory illness, blood pressure, cardiovascular illness, salivary progesterone, psychological well-being, arthritis, diabetes, circulating testosterone, immune function, and oxidative stress. While a strong relationship between facial adiposity, attractiveness, perceived health and actual health outcomes has been reported, there are a few limitations to the current evidence presented in favor of facial adiposity as an important contributor to health and attractiveness judgments. It is noteworthy that the majority of the studies published on facial adiposity used Caucasian students in their early to mid-twenties as both stimuli and raters, with isolated studies being done on Asian, Hispanic and African faces. A study by Batres and Perrett (2014) also found that people without internet access preferred female faces with higher levels of adiposity compared to people with internet access. This finding highlights a key shortcoming in the field, as most current published studies may provide a biased or incomplete picture of the way in which attractiveness and health judgments are made from facial cues related to body weight. It is therefore important for researchers to expand the diversity of sampled populations in future studies, as environmental pressures, and media exposure to Western weight ideals may differ substantially from region to region or culture to culture. It is also recommended that studies take socio-economic status of participants into account if possible.

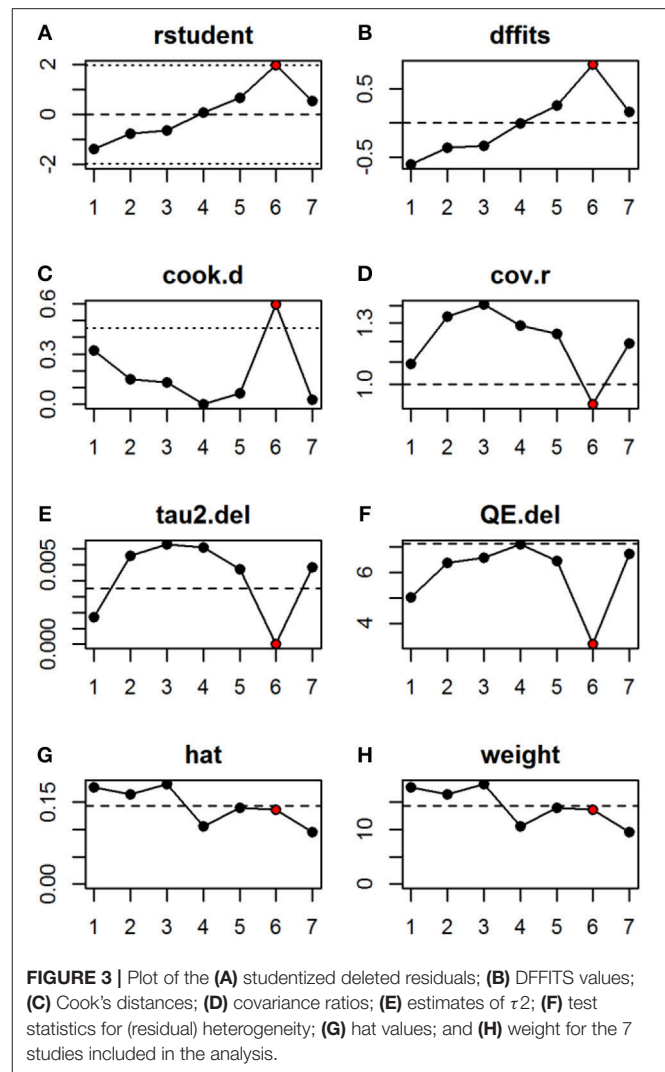


FIGURE 3 | Plot of the (A) studentized deleted residuals; (B) DFFITS values; (C) Cook's distances; (D) covariance ratios; (E) estimates of τ^2 ; (F) test statistics for (residual) heterogeneity; (G) hat values; and (H) weight for the 7 studies included in the analysis.

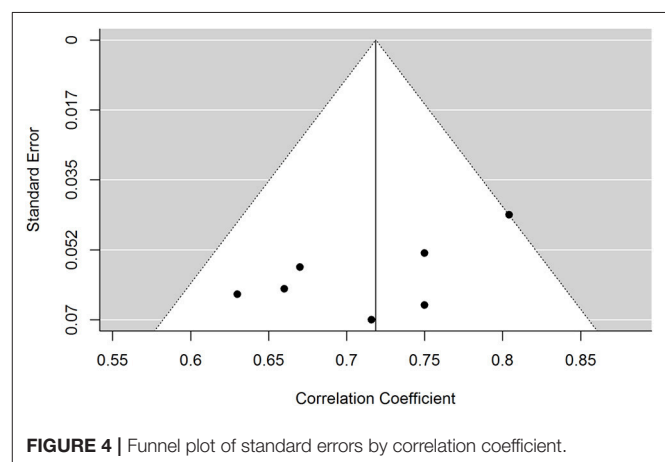


FIGURE 4 | Funnel plot of standard errors by correlation coefficient.

There also appears to be important differences in judgments made by males and females regarding adiposity as a cue to health and attractiveness. For example, a link has been found

between adiposity and immune responsiveness in male Rantala et al., 2013a, but not female faces Rantala et al., 2013b. There is also at least some evidence to suggest that Western body ideals may have a differential effect on males and females, where ideal attractiveness is associated with a lower BMI compared to ideal weight for females (Coetzee et al., 2011). Due to the fact that research in the field is highly reliant on cross-gender judgments of health and attractiveness, most studies would be well-advised to be sensitive to the nuanced differences in how people of opposite genders judge one another based on facial cues related to adiposity, especially across socio-economic or cultural boundaries.

In addition to the socio-economic, cultural and gender factors, researchers also need to pay close attention to stimuli generation and study designs when conducting research. The recent development of 3D facial imaging technology allows researchers to produce 3D facial images that more closely approximate the way people would be able to view another person's face in reality. This is especially important when participants are required to make judgments about facial shape, as 2D facial images contain less overall facial shape information compared to 3D images. Researcher also need to be mindful of the potential drawbacks of using traditional statistical procedures and inadequate sample sizes when the research designs make use of nested structures where multiple participants judge multiple facial stimuli along multiple dimensions.

Despite the limitations presented in this paper, the field of facial morphometrics is a burgeoning field that holds the potential for the development of reliable, inexpensive and non-invasive methods for disease detection and monitoring. For example, the Andreu et al., (2016) is a personal health monitor that integrates a 3D optical scanner, multispectral cameras and gas detection sensor for collecting data of individuals who stand in front of the mirror. One of the key features of the Wize Mirror is the use of facial morphometric analysis to predict cardio-metabolic risk. Studies by Kocabey et al. (2017) and Barr et al. (2018) also reported on the development of Face-to-BMI- Systems that predicts BMI from facial images found

on social media platforms, for example. The results reported by these studies indicate that these computer models could predict actual BMI from facial cues alone to a degree of accuracy very similar to that of human observers. With rapid advancement of these types of technologies, exciting possibilities begin to open up for continuous monitoring of health risk associated with BMI.

As mentioned throughout this paper, one of the key conditions that need to be met for facial adiposity to serve as a valid cue to attractiveness and health, is that people need to be able to accurately judge body mass from facial cues. The meta-analysis presented in this review found that participants can reliably estimate BMI from facial cues alone ($r = 0.71$, $n = 458$) and that ethnicity of the rated faces did not mediate the relationship between perceived facial adiposity and estimates of body mass. This is an important finding as it demonstrates that cues related to facial adiposity can be reliably detected by participants and also reliably used to make inferences about another person's body weight. Due to the strong relationship between body weight and negative health outcomes, accurate judgment of facial cues related to body weight, is a key factor in allowing us to make inferences about another person's health.

AUTHOR CONTRIBUTIONS

SdJ and VC contributed conception and design of the study; SdJ and VC organized the database; SdJ performed the statistical analysis; SdJ and VC wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

The preparation of this paper was financially supported by a Department of Higher Education and Training (DHET) Research Development Grant and a National Research Foundation (NRF) Competitive Programme for Rated Researchers (99078).

REFERENCES

- Andreu, Y., Chiarugi, F., Colantonio, S., Giannakakis, G., Giorgi, D., Henriquez, P., et al. (2016). Wize Mirror - a smart, multisensory cardio-metabolic risk monitoring system. *Comput. Vis. Image Underst.* 148, 3–22. doi: 10.1016/j.cviu.2016.03.018
- Axelsson, J., Sundelin, T., Olsson, M. J., Sorjonen, K., Axelsson, C., Lasselin, J., et al. (2018). Identification of acutely sick people and facial cues of sickness. *Proc. Biol. Sci.* 285:20172430. doi: 10.1098/rspb.2017.2430
- Barr, M., Guo, G., Colby, S., Olfert, M., Barr, M. L., Guo, G., et al. (2018). Detecting body mass index from a facial photograph in lifestyle intervention. *Technologies* 6:83. doi: 10.3390/technologies6030083
- Batres, C., Kannan, M., and Perrett, D. I. (2017). Familiarity with own population's appearance influences facial preferences. *Hum. Nat. Interdiscip. Biosoc. Perspect.* 28, 344–354. doi: 10.1007/s12110-017-9289-8
- Batres, C., and Perrett, D. I. (2014). The influence of the digital divide on face preferences in el salvador: people without internet access prefer more feminine men, more masculine women, and women with higher adiposity. *PLoS ONE* 9:e100966. doi: 10.1371/journal.pone.0100966
- Batres, C., and Perrett, D. I. (2017). How the harsh environment of an army training camp changes human (*Homo sapiens*) facial preferences. *Ethology* 123, 61–68. doi: 10.1111/eth.12571
- Boothroyd, L. G., Scott, I., Gray, A. W., Coombes, C. I., and Pound, N. (2013). Male facial masculinity as a cue to health outcomes. *Evol. Psychol.* 11, 1044–1058. doi: 10.1177/147470491301100508
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* 1, 97–111. doi: 10.1002/jrsm.12
- Cai, Z., Hahn, A. C., Zhang, W., Holzleitner, I. J., Lee, A. J., DeBruine, L. M., et al., (in press). No evidence that facial attractiveness, femininity, averageness, or coloration are cues to susceptibility to infectious illnesses in a university sample of young adult women. *Evol. Hum. Behav.* doi: 10.1016/j.evolhumbehav.2018.10.002
- Coetzee, V. (2011). *The Effect of Weight on Health and Face Perception: A Cross-Cultural Perspective*. Available online at: <https://search.proquest.com/docview/1512387626?accountid=14717>
- Coetzee, V., Chen, J., Perrett, D. I., and Stephen, I. D. (2010). Deciphering faces: quantifiable visual cues to weight. *Perception* 39, 51–61. doi: 10.1068/p6560

- Coetzee, V., Faerber, S. J., Greeff, J. M., Lefevre, C. E., Re, D. E., and Perrett, D. I. (2012). African perceptions of female attractiveness. *PLoS ONE* 7:e48116. doi: 10.1371/journal.pone.0048116
- Coetzee, V., Greeff, J. M., Stephen, I. D., and Perrett, D. I. (2014). Cross-cultural agreement in facial attractiveness preferences: the role of ethnicity and gender. *PLoS ONE* 9:e99629. doi: 10.1371/journal.pone.0099629
- Coetzee, V., Perrett, D. I., and Stephen, I. D. (2009). Facial adiposity: a cue to health? *Perception* 38, 1700–1711. doi: 10.1068/p6423
- Coetzee, V., Re, D., Perrett, D. I., Tiddeman, B. P., and Xiao, D. (2011). Judging the health and attractiveness of female faces: is the most attractive level of facial adiposity also considered the healthiest? *Body Image* 8, 190–193. doi: 10.1016/j.bodyim.2010.12.003
- Del Re, A., and Hoyt, W. (2012). *MAC: Meta-Analysis With Correlations. R package version 1.1*. Available online at: <https://CRAN.R-project.org/package=MAC>
- Després, J. P., and Lemieux, I. (2006). Abdominal obesity and metabolic syndrome. *Nature* 444, 881–887. doi: 10.1038/nature05488
- Dixon, J. B. (2010). The effect of obesity on health outcomes. *Mol. Cell. Endocrinol.* 316, 104–108. doi: 10.1016/j.mce.2009.07.008
- Dixon, B. J., Li, B., and Dixon, A. F. (2010a). Female waist-to-hip ratio, body mass index and sexual attractiveness in China. *Curr. Zool.* 56, 175–181.
- Dixon, B. J., Little, A. C., Dixon, H. G., and Brooks, R. C. (2017). Do prevailing environmental factors influence human preferences for facial morphology? *Behav. Ecol.* 28, 1217–1227. doi: 10.1093/beheco/axx067
- Dixon, B. J., Sagata, K., Linklater, W. L., and Dixon, A. F. (2010b). Male preferences for female waist-to-hip ratio and body mass index in the highlands of Papua New Guinea. *Am. J. Phys. Anthropol.* 141, 620–625. doi: 10.1002/ajpa.21181
- Dixon, B. J., Vasey, P. L., Sagata, K., Sibanda, N., Linklater, W. L., and Dixon, A. F. (2011). Men's preferences for women's breast morphology in New Zealand, Samoa, and Papua New Guinea. *Arch. Sex. Behav.* 40, 1271–1279. doi: 10.1007/s10508-010-9680-6
- Dobner, J., and Kaser, S. (2018). Body mass index and the risk of infection - from underweight to obesity. *Clin. Microbiol. Infect.* 24, 24–28. doi: 10.1016/j.cmi.2017.02.013
- Fisher, C. I., Fincher, C. L., Hahn, A. C., DeBruine, L. M., and Jones, B. C. (2013). Individual differences in pathogen disgust predict men's, but not women's, preferences for facial cues of weight. *Personal. Individ. Differ.* 55, 860–863. doi: 10.1016/j.paid.2013.07.015
- Fisher, C. I., Fincher, C. L., Hahn, A. C., Little, A. C., DeBruine, L. M., and Jones, B. C. (2014a). Do assortative preferences contribute to assortative mating for adiposity? *Br. J. Psychol.* 105, 474–485. doi: 10.1111/bjop.12055
- Fisher, C. I., Hahn, A. C., DeBruine, L. M., and Jones, B. C. (2014b). Integrating shape cues of adiposity and color information when judging facial health and attractiveness. *Perception* 43, 499–508. doi: 10.1068/p7728
- Folstad, I., and Karter, A. (1992). Parasites, bright males, and the immunocompetence handicap. *Am. Nat.* 139, 603–622. doi: 10.1086/285346
- Foo, Y. Z., Simmons, L. W., and Rhodes, G. (2017). Predictors of facial attractiveness and health in humans. *Sci. Rep. Nat. Publ. Group* 7:39731. doi: 10.1038/srep39731
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., and McCormick, C. M. (2015). Evidence from Meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS ONE* 10:e0132726. doi: 10.1371/journal.pone.0132726
- Gouda-Vossos, A., Nakagawa, S., Dixon, B. J., and Brooks, R. C. (2018). Mate choice copying in humans: a systematic review and meta-analysis. *Adapt. Hum. Behav. Physiol.* 4, 364–386. doi: 10.1007/s40750-018-0099-y
- Grammer, K., and Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness. *J. Comp. Psychol.* 108, 233–242. doi: 10.1037/0735-7036.108.3.233
- Groesz, L. M., Levine, M. P., and Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: a meta-analytic review. *Int. J. Eat. Disord.* 31, 1–16. doi: 10.1002/eat.10005
- Gurevitch, J., Koricheva, J., Nakagawa, S., and Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature* 555, 175–182. doi: 10.1038/nature25753
- Hamilton, W. D., and Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites? *Science* 218, 384–387. doi: 10.1126/science.7123238
- Han, C., Hahn, A. C., Fisher, C. I., DeBruine, L. M., and Jones, B. C. (2016). Women's facial attractiveness is related to their body mass index but not their salivary cortisol. *Am. J. Hum. Biol.* 28, 352–355. doi: 10.1002/ajhb.22792
- Harrison, K., and Hefner, V. (2006). Media exposure, current and future body ideals, and disordered eating among preadolescent girls: a longitudinal panel study. *J. Youth Adolesc.* 35, 146–156. doi: 10.1007/s10964-005-9008-3
- Henderson, A. J., Holzleitner, I. J., Talamas, S. N., and Perrett, D. I. (2016). Perception of health from facial cues. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371:20150380. doi: 10.1098/rstb.2015.0380
- Henderson, J. J. A., and Anglin, J. M. (2003). Facial attractiveness predicts longevity. *Evol. Hum. Behav.* 24, 351–356. doi: 10.1016/S1090-5138(03)00036-9
- Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ* 327, 557–560. doi: 10.1136/bmj.327.7414.557
- Hunter, J., and Schmidt, F. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. 2nd Edn*. Thousand Oaks, CA: Sage.
- Jones, A. L., Kramer, R. S. S., and Ward, R. (2012). Signals of personality and health: the contributions of facial shape, skin texture, and viewing angle. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1353–1361. doi: 10.1037/a0027078
- Jones, B. C., Fincher, C. L., Welling, L. L. M., Little, A. C., Feinberg, D. R., Watkins, C. D., et al. (2013). Salivary cortisol and pathogen disgust predict men's preferences for feminine shape cues in women's faces. *Biol. Psychol.* 92, 233–240. doi: 10.1016/j.biopsycho.2012.11.014
- Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception* 33, 569–576. doi: 10.1068/p3463
- Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B. P., Burt, D. M., and Perrett, D. I. (2001). Facial symmetry and judgements of apparent health: Support for a “good genes” explanation of the attractiveness–symmetry relationship. *Evol. Hum. Behav.* 22, 417–429. doi: 10.1016/S1090-5138(01)00083-6
- Kalick, S. M., Zebrowitz, L. A., Langlois, J. H., and Johnson, R. M. (1998). Does human facial attractiveness honestly advertise health? longitudinal data on an evolutionary question. *Psychol. Sci.* 9, 8–13. doi: 10.1111/1467-9280.00002
- Klaczynski, P., Daniel, D. B., and Keller, P. S. (2009). Appearance idealization, body esteem, causal attributions, and ethnic variations in the development of obesity stereotypes. *J. Appl. Dev. Psychol.* 30, 537–551. doi: 10.1016/j.appdev.2008.12.031
- Kocabay, E., Camurcu, M., Ofli, F., Aytaç, Y., Marin, J., Torralba, A., et al. (2017). Face-to-BMI: using computer vision to infer body mass index on social media. *arXiv:1703.03156v1 [cs.HC]*.
- Kopelman, P. (2007). Health risks associated with overweight and obesity. *Obes. Rev.* 8, 13–17. doi: 10.1111/j.1467-789X.2007.00311.x
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* 126, 390–423. doi: 10.1037/0033-2909.126.3.390
- Lee, B. J., and Kim, J. Y. (2014). Predicting visceral obesity based on facial characteristics. *BMC Complement. Altern. Med.* 14:248. doi: 10.1186/1472-6882-14-248
- Levine, J. A., Ray, A., and Jensen, M. D. (1998). Relation between chubby cheeks and visceral fat. *N. Engl. J. Med.* 339, 1946–1947. doi: 10.1056/NEJM199812243392619
- Little, A. C., Burriss, R. P., Jones, B. C., and Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evol. Hum. Behav.* 28, 18–27. doi: 10.1016/j.evolhumbehav.2006.09.002
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Martinson, M. L., and Vasunilashorn, S. M. (2016). The long-arm of adolescent weight status on later life depressive symptoms. *Age Ageing* 45, 389–395. doi: 10.1093/ageing/afw020
- Mayer, C., Windhager, S., Schaefer, K., and Mitteroecker, P. (2017). BMI and WHR are reflected in female facial shape and texture: a geometric morphometric image analysis. *PLoS ONE* 12:e0169336. doi: 10.1371/journal.pone.0169336

- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed.1000097
- Nowak, J., Pawłowski, B., Borkowska, B., Augustyniak, D., and Drulis-Kawa, Z. (2018). No evidence for the immunocompetence handicap hypothesis in male humans. *Sci. Rep.* 8:7392. doi: 10.1038/s41598-018-25694-0
- Park, J. H., van Leeuwen, F., and Stephen, I. D. (2012). Homeliness is in the disgust sensitivity of the beholder: relatively unattractive faces appear especially unattractive to individuals higher in pathogen disgust. *Evol. Hum. Behav.* 33, 569–577. doi: 10.1016/j.evolhumbehav.2012.02.005
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., and Edwards, R. (1999). Symmetry and Human Facial Attractiveness. *Evol. Hum. Behav.* 20, 295–307. doi: 10.1016/S1090-5138(99)00014-8
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Phalane, K. G., Tribe, C., Steel, H. C., Cholo, M. C., and Coetzee, V. (2017). Facial appearance reveals immunity in African men. *Sci. Rep. Nat. Publ. Group* 7, 1–9. doi: 10.1038/s41598-017-08015-9
- Quintana, D. S. (2015). From pre-registration to publication: a non-technical primer for conducting a meta-analysis to synthesize correlational data. *Front. Psychol.* 6:1549. doi: 10.3389/fpsyg.2015.01549
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rantala, M. J., Coetzee, V., Moore, F. R., Skrinda, I., Kecko, S., Krama, T., et al., (2013a). Adiposity, compared with masculinity, serves as a more valid cue to immunocompetence in human mate choice. *Proc. Biol. Sci.* 280:20122495. doi: 10.1098/rspb.2012.2495
- Rantala, M. J., Coetzee, V., Moore, F. R., Skrinda, I., Kecko, S., Krama, T., et al., (2013b). Facial attractiveness is related to women's cortisol and body fat, but not with immune responsiveness. *Biol. Lett.* 9:20130255. doi: 10.1098/rsbl.2013.0255
- Re, D. E., Coetzee, V., Xiao, D., Bult, D., Tiddeman, B. P., Boothroyd, L. G., et al. (2011). Viewing heavy bodies enhances preferences for facial adiposity. *J. Evol. Psychol.* 9, 295–308. doi: 10.1556/JEP.9.2011.4.2
- Re, D. E., Holzleitner, I. J., Coetzee, V., Tiddeman, B. P., Xiao, D. K., and Perrett, D. I. (2013). Wie viel gewicht liegt im gesicht? *J. Für Ästhetische Chir.* 6, 141–146. doi: 10.1007/s12631-013-0240-5
- Re, D. E., and Perrett, D. I. (2014). The effects of facial adiposity on attractiveness and perceived leadership ability. *Q. J. Exp. Psychol.* 2006 67, 676–686. doi: 10.1080/17470218.2013.825635
- Re, D. E., and Rule, N. O. (2016a). "Appearance and physiognomy," in *APA Handbook of Nonverbal Communication*, eds. D. Matsumoto, H. Hwang, and M. Frank (Washington DC: American Psychological Association), 221–256.
- Re, D. E., and Rule, N. O. (2016b). Heavy matters: the relationship between just noticeable differences in perceptions of facial adiposity and facial attractiveness. *Soc. Psychol. Personal. Sci.* 7, 69–76. doi: 10.1177/1948550615599829
- Reither, E. N., Hauser, R. M., and Swallen, K. C. (2009). Predicting adult health and mortality from adolescent facial characteristics in yearbook photographs*. *Demography* 46, 27–41. doi: 10.1353/dem.0.0037
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Rhodes, G., Yoshikawa, S., Palermo, R., Simmons, L. W., Peters, M., Lee, K., et al., (2007). Perceived health contributes to the attractiveness of facial symmetry, averageness, and sexual dimorphism. *Perception* 36, 1244–1252. doi: 10.1068/p5712
- Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., and McKay, R. (2001). Do facial averageness and symmetry signal health? *Evol. Hum. Behav.* 22, 31–46. doi: 10.1016/S1090-5138(00)00060-X
- Ritz, B. W., and Gardner, E. M. (2006). Malnutrition and energy restriction differentially affect viral immunity. *J. Nutr.* 136, 1141–1144. doi: 10.1093/jn/136.5.1141
- Roberts, M. L., Buchanan, K. L., and Evans, M. R. (2004). Testing the immunocompetence handicap hypothesis: a review of the evidence. *Anim. Behav.* 68, 227–239. doi: 10.1016/j.anbehav.2004.05.001
- Russell, R., Porcheron, A., Sweda, J. R., Jones, A. L., Mauger, E., and Morizot, F. (2016). Facial contrast is a cue for perceiving health from the face. *J. Exp. Psychol. Hum. Percept. Perform.* 42:1354. doi: 10.1037/xhp0000219
- Scheib, J. E., Gangestad, S. W., and Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proc. Biol. Sci.* 266, 1913–1917. doi: 10.1098/rspb.1999.0866
- Schneider, T. M., Hecht, H., Stevanov, J., and Carbon, C.-C. (2013). Cross-ethnic assessment of body weight and height on the basis of faces. *Personal. Individ. Differ.* 55, 356–360. doi: 10.1016/j.paid.2013.03.022
- Shackelford, T. K., and Larsen, R. J. (1999). Facial attractiveness and physical health. *Evol. Hum. Behav.* 20, 71–76. doi: 10.1016/S1090-5138(98)00036-1
- Silventoinen, K., Jelenkovic, A., Sund, R., Yokoyama, Y., Hur, Y.-M., Cozen, W., et al. (2017). Differences in genetic and environmental variation in adult BMI by sex, age, time period, and region: an individual-based pooled analysis of 40 twin cohorts. *Am. J. Clin. Nutr.* 106, 457–466. doi: 10.3945/ajcn.117.153643
- Singh, D., Dixon, B. J., Jessop, T. S., Morgan, B., and Dixon, A. F. (2010). Cross-cultural consensus for waist-hip ratio and women's attractiveness. *Evol. Hum. Behav.* 31, 176–181.
- Stephen, I. D., Coetzee, V., and Perrett, D. I. (2011). Carotenoid and melanin pigment coloration affect perceived human health. *Evol. Hum. Behav.* 32, 216–227. doi: 10.1016/j.evolhumbehav.2010.09.003
- Stephen, I. D., Hiew, V., Coetzee, V., Tiddeman, B. P., and Perrett, D. I. (2017). Facial shape analysis identifies valid cues to aspects of physiological health in Caucasian, Asian, and African populations. *Front. Psychol.* 8:1883. doi: 10.3389/fpsyg.2017.01883
- Stephen, I. D., Law Smith, M. J., Stirrat, M. R., and Perrett, D. I. (2009). Facial skin coloration affects perceived health of human faces. *Int. J. Primatol.* 30, 845–857. doi: 10.1007/s10764-009-9380-z
- Swami, V., and Tovée, M. J. (2007). Differences in attractiveness preferences between observers in low- and high-resource environments in Thailand. *J. Evol. Psychol.* 5, 149–160. doi: 10.1556/JEP.2007.1005
- Swami, V., and Tovée, M. J. (2013). Resource security impacts men's female breast size preferences. *PLoS ONE* 8:e57623. doi: 10.1371/journal.pone.0057623
- Talamas, S. N., Mavor, K. I., and Perrett, D. I. (2016). Blinded by beauty: attractiveness bias and accurate perceptions of academic performance. *PLoS ONE* 11: e0148284. doi: 10.1371/journal.pone.0148284
- Thornhill, R., and Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Hum. Nat. Hawthorne N* 4, 237–269. doi: 10.1007/BF02692201
- Thornhill, R., and Gangestad, S. W. (1999). Facial attractiveness. *Trends Cogn. Sci.* 3, 452–460. doi: 10.1016/S1364-6613(99)01403-5
- Thornhill, R., and Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evol. Hum. Behav.* 27, 131–144. doi: 10.1016/j.evolhumbehav.2005.06.001
- Tinlin, R. M., Watkins, C. D., Welling, L. L. M., DeBruine, L. M., Al-Dujaili, E. A. S., and Jones, B. C. (2013). Perceived facial adiposity conveys information about women's health. *Br. J. Psychol. Lond. Engl.* 104, 235–248. doi: 10.1111/j.2044-8295.2012.02117.x
- Tovée, M. J., Furnham, A., and Swami, V. (2007). "Healthy body equals beautiful body? changing perceptions of health and attractiveness with shifting socioeconomic status," in *The Body Beautiful: Evolutionary and Sociocultural Perspectives*, eds. V. Swami and A. Furnham (London: Palgrave Macmillan UK), 108–128.
- Tovée, M. J., Swami, V., Furnham, A., and Mangalparsad, R. (2006). Changing perceptions of attractiveness as observers are exposed to a different culture. *Evol. Hum. Behav.* 27, 443–456. doi: 10.1016/j.evolhumbehav.2006.05.004
- Valavanidis, A., Vlachogianni, T., and Fiotakis, C. (2009). 8-hydroxy-2'-deoxyguanosine (8-OHdG): a critical biomarker of oxidative stress and carcinogenesis. *J. Environ. Sci. Health Part C Environ. Carcinog. Ecotoxicol. Rev.* 27, 120–139.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48. doi: 10.18637/jss.v036.i03
- Viechtbauer, W., and Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Res. Synth. Methods* 1, 112–125. doi: 10.1002/jrsm.11

- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., and Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, “all must have prizes.” *Psychol. Bull.* 122:203.
- Weeden, J., and Sabini, J. (2005). Physical attractiveness and health in Western societies: a review. *Psychol. Bull.* 131, 635–653. doi: 10.1037/0033-2909.131.5.635
- Wen, L., and Guo, G. (2013). A computational approach to body mass index prediction from face images. *Image Vis. Comput.* 31, 392–400. doi: 10.1016/j.imavis.2013.03.001
- Weston, T. D., Hass, N. C., and Lim, S.-L. (2015). The effect of sad facial expressions on weight judgment. *Front. Psychol.* 6. doi: 10.3389/fpsyg.2015.00417
- Windhager, S., Bookstein, F. L., Mueller, H., Zunner, E., Kirchengast, S., and Schaefer, K. (2018). Calibrating facial morphs for use as stimuli in biological studies of social perception. *Sci. Rep. Nat. Publ. Group* 8, 1–9. doi: 10.1038/s41598-018-24911-0
- Windhager, S., Patocka, K., and Schaefer, K. (2013). Body fat and facial shape are correlated in female adolescents. *Am. J. Hum. Biol.* 25, 847–850. doi: 10.1002/ajhb.22444
- Wolffhechel, K., Hahn, A. C., Jarmer, H., Fisher, C. I., Jones, B. C., and DeBruine, L. M. (2015). Testing the utility of a data-driven approach for assessing BMI from face images. *PLoS ONE* 10:e0140347. doi: 10.1371/journal.pone.0140347
- World Health Organization (2018). *Body Mass Index - BMI*. Available online at: <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>
- Zaccardi, F., Dhalwani, N. N., Papamargaritis, D., Webb, D. R., Murphy, G. J., Davies, M. J., et al., (2017). Nonlinear association of BMI with all-cause and cardiovascular mortality in type 2 diabetes mellitus: a systematic review and meta-analysis of 414,587 participants in prospective studies. *Diabetologia* 60, 240–248. doi: 10.1007/s00125-016-4162-6
- Zebrowitz, L. A., Hall, J. A., Murphy, N. A., and Rhodes, G. (2002). Looking smart and looking good: facial cues to intelligence and their origins. *Pers. Soc. Psychol. Bull.* 28, 238–249. doi: 10.1177/0146167202282009

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 de Jager, Coetzee and Coetzee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Influence of Body Composition Effects on Male Facial Masculinity and Attractiveness

Xue Lei^{1*}, Iris J. Holzleitner² and David I. Perrett¹

¹ School of Psychology & Neuroscience, University of St Andrews, St Andrews, United Kingdom, ² Institute of Neuroscience & Psychology, University of Glasgow, Glasgow, United Kingdom

Body mass index (BMI) and its facial correlates influence a range of perceptions including masculinity and attractiveness. BMI conflates body fat and muscle which are sexually dimorphic because men typically have more muscle but less fat than women. We therefore investigated the influence of facial correlates of body composition (fat mass and muscle mass) on the perception of masculinity in male faces. Women have been found to prefer more masculine looking men when considering short-term relationships compared with long-term relationships. We therefore conducted a second study of heterosexual women's preferences for facial correlates of fat and muscle mass under long and short relationship contexts. We digitally transformed face shape simulating the effects of raised and lowered levels of body fat or muscle, controlling for each other, height and age. In Study 1, participants rated masculinity of shape-transformed male faces. The face shape correlates of muscle mass profoundly enhanced perceived masculinity but the face shape correlates of fat mass only affected the perception of masculinity in underweight to low normal weight men. In Study 2, we asked two groups of women to optimize male face images (by adjusting the shape correlates of fat and muscle) to most resemble someone they would prefer, either for a short-term sexual relationship or for a long-term relationship. The results were consistent across the two participant groups: women preferred the appearance of male faces associated with a higher muscle mass for short-term compared with long-term relationships. No difference was found in women's preference for the face shape correlates of fat mass between the two relationship contexts. These findings suggest that the facial correlates of body fat and muscle have distinct impacts on the perception of male masculinity and on women's preferences. The findings indicate that body composition needs to be taken into consideration in psychological studies involving body weight.

Keywords: body composition, fat, muscle, masculinity, face preference, short-term relationship, long-term relationship, relationship context

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Karel Kleisner,
Charles University, Czechia
Thomas Richardson,
University of Manchester,
United Kingdom

*Correspondence:

Xue Lei
xl55@st-andrews.ac.uk

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 October 2018

Accepted: 10 December 2018

Published: 04 January 2019

Citation:

Lei X, Holzleitner IJ and Perrett DI
(2019) The Influence of Body
Composition Effects on Male Facial
Masculinity and Attractiveness.
Front. Psychol. 9:2658.
doi: 10.3389/fpsyg.2018.02658

INTRODUCTION

Research on women's preference for male facial masculinity over the past two decades is marked by inconsistent findings. Some studies found that masculine faces were preferred by women (e.g., Rhodes et al., 2003; DeBruine et al., 2006; Feinberg et al., 2008; Little et al., 2008; Saxton et al., 2009; Jones et al., 2018), whereas other studies have reported a preference for femininity in men (e.g., Perrett et al., 1998; Penton-Voak et al., 1999; Little et al., 2002; Scott et al., 2010), and yet

other studies report no overall preference for sexual dimorphism (e.g., Swaddle and Reiersen, 2002; Cornwell et al., 2004).

Variability in methods has been proposed to account for the differences in results (Rhodes, 2006), yet by directly comparing commonly used methods to measure women's preferences for male facial masculinity, DeBruine et al. (2006) found that different methods can produce similar results. Alternatively, individual differences in self-rated attractiveness, relationship status, own-health condition, exposure to violence, pathogen disgust sensitivity and resource availability might contribute to the variation in results (Holzleitner and Perrett, 2017). One factor that has been found to have a consistent effect on women's preference for male masculinity is relationship context. Using computer graphics techniques to manipulate masculinity in male facial shape, women show a stronger preference for facial masculinity when choosing short-term partners compared to long-term partners (Little et al., 2002; Penton-Voak et al., 2003; Jones et al., 2018). In addition, this relationship context effect was more pronounced in women with partners and not found in those taking hormonal contraception pills (Little et al., 2002). This preference for masculinity in men as short-term partners has been found with a range of stimuli and modalities, including face, body, voice, and odor (Little et al., 2011a).

Sexual Strategies theory proposes that females have evolved distinct strategies to solve different problems they may encounter when pursuing a short-term or long-term relationship (Buss and Schmitt, 1993). As women's reproductive success is restricted by the resources and protection they can obtain from men, women should prefer long-term partners who are more likely to provide paternal care, reliable resources and protection. Masculinity is perceptually associated with some negative personality traits, which might explain why women prefer less masculine men for long-term partners. Indeed, perceived facial masculinity was found to increase perceived dominance (Boothroyd et al., 2007), lower perceived paternal investment (Boothroyd et al., 2007) and decrease perceived trustworthiness (Perrett et al., 1998). Complementing these findings, several studies have found that high testosterone (an androgen contributing to male sexual dimorphism) is associated with lower likelihood of marriage, higher divorce rates and higher rates of domestic disputes (Julian and McKenry, 1989; Booth and Dabbs, 1993; Booth et al., 2000). Hence, less masculine men may be advantageous for long-term relationships.

In short-term relationships, women need not be restricted by consideration of paternal investment. Therefore, selection of partners may be guided by cues to long-term health and 'good genes' for immunity against currently prevalent pathogens that can be passed on to offspring (Gangestad et al., 2005). Masculinity is argued to be one cue to good genes as part of the immunocompetence handicap hypothesis (Folstad and Karter, 1992). This hypothesis states that testosterone has an immunosuppressive effect. Masculine men need a strong immune system to resist the immunosuppressive effect. Masculinity may therefore signal a strong immune system in men. Although studies examining the relationship between testosterone and immune function have produced mixed results, a recent cross-species meta-analysis revealed a medium-sized effect from

experimental studies which elevate testosterone artificially and find a concomitant decline in immune function (Foo et al., 2017).

While a considerable number of studies have focused on the role of testosterone in suppressing immune function, it is relevant that testosterone has also been found to play a key role in maintaining men's cardiovascular health. A deficiency in testosterone is associated with increased central adiposity, reduced insulin sensitivity, impaired glucose tolerance and increased cholesterol, which are all found in metabolic syndrome and type 2 diabetes and are detrimental to cardiovascular health (Kelly and Jones, 2013). Although there is debate about whether lower levels of testosterone cause cardiovascular diseases directly or whether decreased testosterone is a by-product of poor health, clinical studies have found that testosterone replacement therapy is effective in improving health in metabolic syndromes (Elagizi et al., 2018). If masculinity is heritable, masculinity may be a cue to current health and to genes for good health.

Despite the prolific research on the effect of masculine traits (e.g., faces, voices, odors) on attractiveness, few studies have explored the role that muscle plays. This is surprising considering the fact that higher muscle mass to lower fat mass is a typical masculine feature in humans (Wells, 2007) because testosterone promotes both muscle and bone growth (Mooradian et al., 1987). Thus, measures of muscle might be strong cues to masculinity. It follows that one may expect men with high muscle to be preferred by women, especially for short-term relationships, as women prefer more masculine looking men for short-term relationships. Indeed, muscular men were found to be preferred by women and have greater mating success (Frederick and Haselton, 2007).

Besides the close relationship between testosterone and muscle mass, muscularity may influence masculinity perception through its association with body size, which is also sexually dimorphic. Men on average are heavier compared to women. Indeed the faces of men with higher body mass index (BMI; weight scaled by the square of height) are perceived as more masculine than men with low BMI (Holzleitner et al., 2014). Therefore, muscular men may be perceived as masculine because they have greater weight. Since body weight is mainly composed of fat and muscle, it raises the question as to whether or not fat mass has a similar effect to muscle mass on male masculinity and attractiveness.

To our knowledge, only one study has explored the role of body composition on the perception of attractiveness in male bodies (Brierley et al., 2016). The results from this study suggest that men with levels of body fat and muscle mass in the healthy BMI range are most preferred by women. This study did not investigate the context of the attractiveness judgments. More importantly, no study has tested the effects of facial correlates of body composition (fat and muscle) on the perception of masculinity and facial attractiveness. Humans rely more heavily on facial attractiveness than physical (body) attractiveness when choosing mates (Currie and Little, 2009). In fact, when given the choice, women gave priority to men's faces over bodies when judging dating partners for both short- and long-term relationships (Confer et al., 2010). These findings highlight the importance of investigating the effect of the facial cues to body composition on attractiveness.

In the current studies, we examine (a) the impact of facial correlates of body composition (fat and muscle) on perceived male facial masculinity, and (b) how the facial correlates of body composition influence women's preference for male faces under short-term and long-term relationship contexts.

Considering that testosterone encourages the growth of muscle, we predict that the facial correlate of muscle mass will be positively correlated with perceived facial masculinity (Hypothesis 1). Since men are heavier than women, a heavier body no matter whether the weight is due to fat mass or muscle mass may lead to higher perceived masculinity. We thus predict the facial correlate of fat mass should also contribute positively to the perception of male facial masculinity (Hypothesis 2). Nevertheless, we expect the face shape correlate of muscle to have a larger effect on perceived facial masculinity than the face shape correlate of fat based on the stronger association between muscle and testosterone than the association between fat and testosterone (Hypothesis 3).

Regarding facial preferences, we predict that women should show a stronger preference for facial cues to increased muscle mass under a short-term relationship context compared to a long-term relationship context (Hypothesis 4). Similarly, we predict a stronger preference for facial cues to increased fat mass in short-term relationships compared to long-term relationships (Hypothesis 5). We also predict that the relationship context effect on preferences will be more apparent for the facial correlates of muscle than the facial correlates of fat (Hypothesis 6). These hypotheses about preferences follow from Hypotheses 1–3 since higher weight, particularly from muscle is expected to increase masculinity.

MATERIALS AND METHODS

Stimuli

To examine the generalizability of findings, we included three sets of faces. One set of three-dimensional (3D) face stimuli, collected using a 3D camera and delineated with 49 landmarks using MorphAnalyser software that included scans of 50 Caucasian men ($M_{\text{age}} \pm SD = 21.2 \pm 2.5$ years, see Holzleitner and Perrett, 2016). A second set of two-dimensional (2D) images matched to the 3D scans were also available for the same 50 men (hereafter referred to as the 2D version of 3D face set). These 2D images were captured under a constant lighting condition using a Fujifilm FinePix S5Pro digital SLR camera (60 mm fixed length lens) in a booth painted with standard white paint. Facial images were captured in full color with participants' hair pulled back. Participants, seated at a set distance from the camera and the same relative eye height to the camera, were asked to maintain a neutral expression. Faces were delineated in PsychoMorph¹ with 189 landmarks and aligned on the left and right pupils (Tiddeman et al., 2001).

A further independent set of 2D face images was collected from 101 Caucasian male participants ($M_{\text{age}} \pm SD = 21.44 \pm 3.33$ years) who were recruited from the University of St Andrews. The

participants contributing to the 3D face set and matched 2D face set did not contribute to the independent 2D face set.

Anthropometric Measurements

Anthropometric data were acquired after removing excess clothing and footwear. Each individual's height was measured with a tape measure (stadiometer), and body composition was measured barefoot using an electrical impedance scale (Tanita SC-330 body composition analyzer), which estimates weight, BMI, fat mass and muscle mass (lean fat-free mass). These estimations take into account information about athletic training (>10 h/week) and norms for each gender. The indicator 'muscle mass' refers to an estimate of the weight of fat-free mass excluding bone mass, and includes contributions from skeletal muscles, smooth muscles and cardiac muscles.

Face Transformation

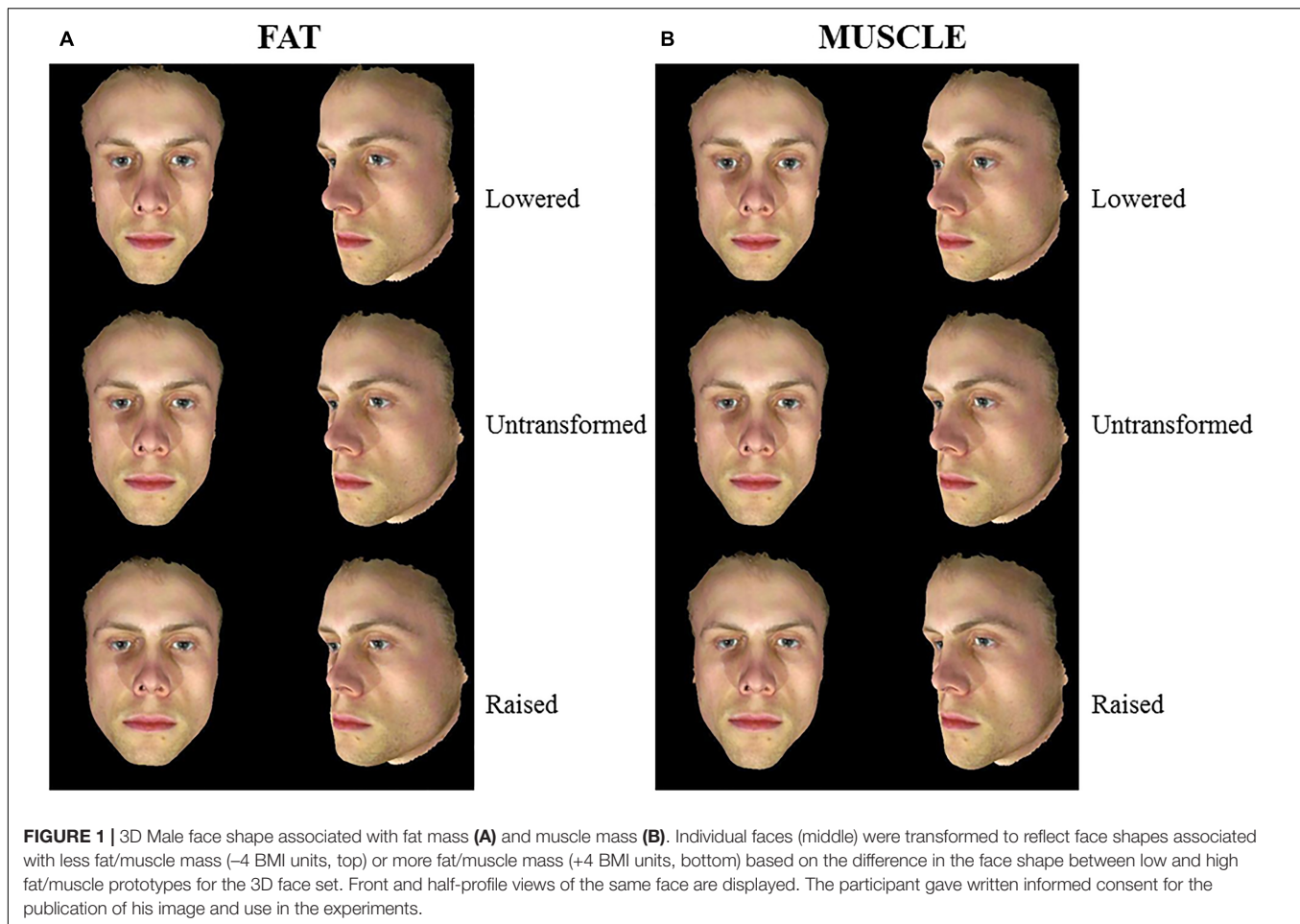
The method used to transform the face shape involves defining the difference in face shape between two groups of faces differentiated along one dimension (e.g., high/low BMI, see Holzleitner and Perrett, 2016; Batres and Perrett, 2017). The difference is then applied to individual face images.

Prototypes associated with high or low fat mass or muscle mass were first created separately for 2D and 3D faces. Prototypes were made by averaging together the nine faces for 3D face set (and matched 2D version of 3D face set) ranked the highest and lowest on the fat mass or muscle mass dimension. This allows a direct comparison between 2D and 3D faces. Since larger individuals usually have higher absolute fat mass and muscle mass than smaller individuals, fat prototypes were created with age, height and muscle mass controlled. Similarly, muscle prototypes were created with age, height and fat mass controlled. Therefore, prototypes differed only in either fat or muscle mass dimension but not in both dimensions (see **Supplementary Material Table S1** for details). Similarly, we created prototypes from the 10 faces ranked highest and lowest in fat or muscle mass dimension for the independent 2D face set.

The fat and muscle prototypes were then used to create shape transforms of five Caucasian male faces. Face shapes were transformed to visualize body composition (fat/muscle mass) differences by adding or subtracting a proportion of the facial shape differences between low and high fat/muscle prototypes. To make the fat- and muscle- transformed images comparable, facial shapes were transformed to the same magnitude in terms of BMI (± 4 BMI units) in 15 steps. This process created three sets of transformed images (using 3D prototypes, 2D version of 3D prototypes and an independent set of 2D prototypes). Each set of transformed images consisted of five identities transformed to lose/gain fat/muscle mass (**Figures 1–3**). For 3D images, both the front view and the half-profile view were created in the transformation process. These two views were combined in one image (**Figure 1**).

All images were masked with the black background to display only the face and neck and to remove confounds arising from hair (DeBruine et al., 2010). 2D images were aligned to have the same pupil positions and resized to 500 × 500 pixels.

¹<http://users.aber.ac.uk/bpt/jpsycomorph/>



STUDY 1: FACIAL CORRELATES OF BODY COMPOSITION AND PERCEIVED MASCULINITY

This study aimed at testing whether facial correlates of body composition (fat mass and muscle mass) influence perceived facial masculinity in males. We tested the following hypotheses:

- (1) Faces associated with more muscle mass will be perceived as more masculine.
- (2) Faces associated with more fat mass will be perceived as more masculine.
- (3) The facial correlate of muscle mass has a larger impact on perceived facial masculinity than the facial correlate of fat mass.

Methods

Ethical approval was received from University of St Andrews Ethics Committee (PS13092). Participants gave written informed consent to perform the experimental tasks.

Participants

Sixty-seven students from the University of St Andrews ($M_{\text{age}} \pm SD = 19.37 \pm 3.84$ years, range 18–45) including

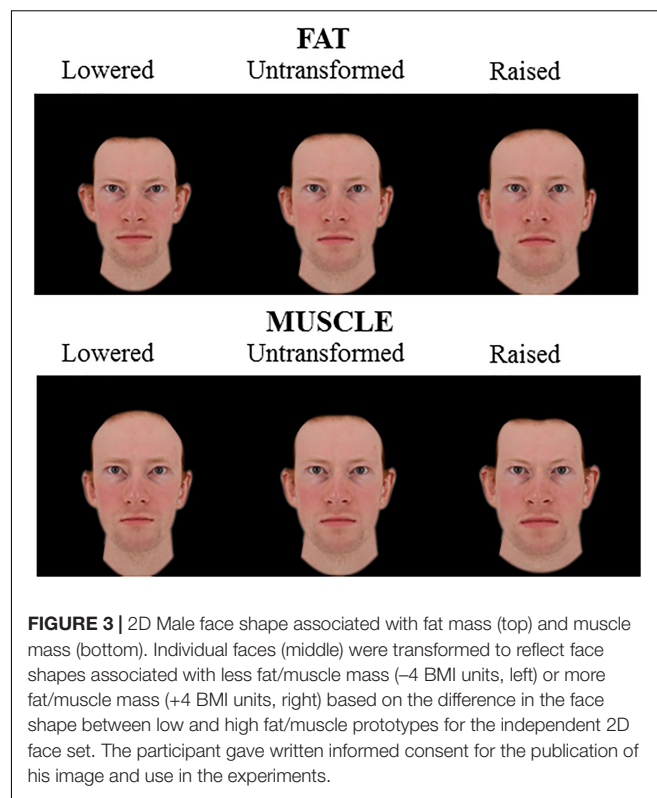
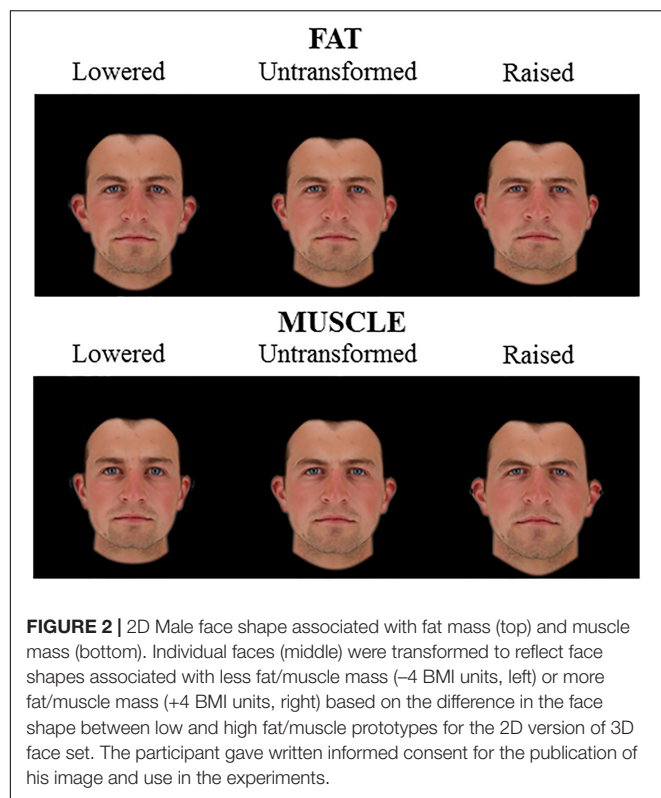
56 females and 9 males (demographics were omitted by two participants; 51 Caucasian) completed this study.

Materials

Stimuli consisted of three face identities transformed to four levels (−4 BMI units, −2.3 BMI units, +2.3 BMI units, +4 BMI units) plus the untransformed image (+0 BMI units). Therefore, there was a total of 81 stimuli: 3 (face identities) \times 3 (face sets: 3D face set, matched 2D version of 3D face set, independent 2D face set) \times 9 [4 BMI levels \times 2 dimensions (fat and muscle) + original face].

Procedure

Participants were asked to complete a demographic questionnaire (age, sex, ethnicity, and sexual orientation). Then faces were presented one at a time in three blocks (each block consisted of a set of faces with muscle and fat transform). Both the order of the trials within blocks and the three blocks were completely randomized. Participants were asked to rate the masculinity (“Please indicate how masculine you perceive this man to be”) of each stimulus face by dragging the cursor on a sliding bar with anchors (1 = least masculine and 7 = most masculine). The starting point of the cursor along the



bar was randomized. There was no time limit to make judgments. The next face was shown only after the participant had adjusted the slider and clicked for the next trial.

Statistical Analysis

For each stimulus type, the mean ratings were calculated across face identities for each participant. The consolidated data were further analyzed in SPSS 24.0 three-way analysis of variance (ANOVA) was run, with the transform dimension (fat/muscle) and the transform level (five levels: −4 BMI units, −2.3 BMI units, no change, +2.3 BMI units, +4 BMI units) included as the independent variables. Face set (three sets) was included as an additional independent variable to

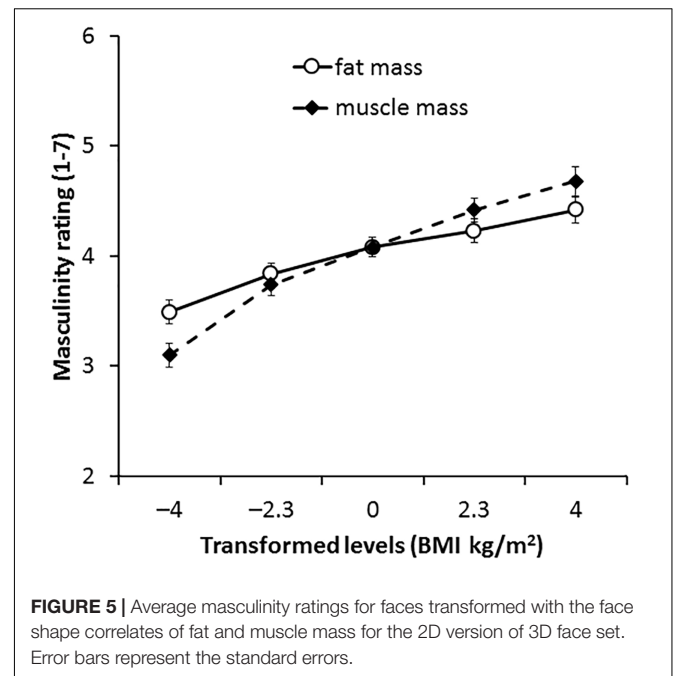
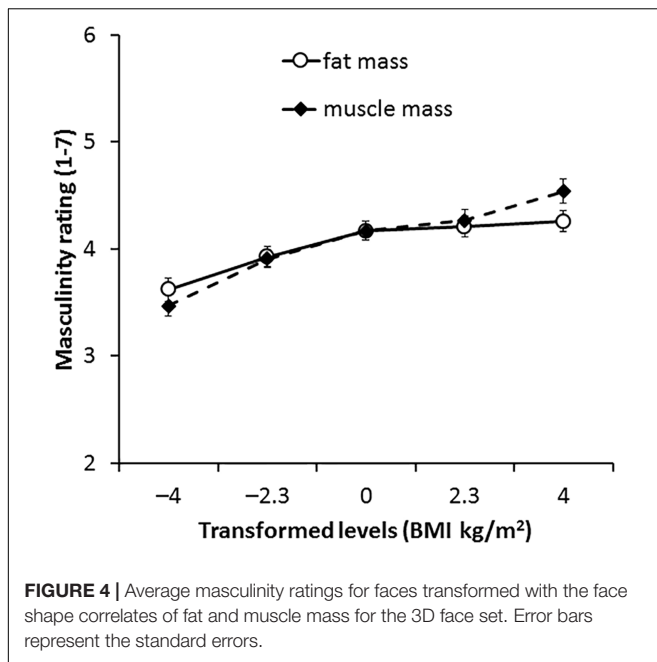
determine if results were consistent across the different samples of faces.

Results

A three-way ANOVA was run to test the transformation attributions made to fat and muscle mass across the three face sets. The results showed non-significant main effects of the transform dimension [$F(1,66) = 0.44, p = 0.507, \eta^2 = 0.007$] and face sets [$F(2,132) = 0.94, p = 0.392, \eta^2 = 0.014$] on masculinity rating, but a significant main effect of transform level [$F(4,264) = 74.80, p < 0.001, \eta^2 = 0.531$] (see **Table 1**). As face shape simulated heavier individuals (higher BMI), the masculinity ratings increased. The interaction between transform dimension and face set was not significant [$F(2,132) = 0.41,$

TABLE 1 | Descriptive statistics of mean masculinity ratings (1–7) (SD) for three sets of faces transformed in fat mass and muscle mass dimensions at five BMI levels.

	−4 BMI	−2.3 BMI	0	+2.3 BMI	+4 BMI
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Fat					
3D face set	3.62 (0.92)	3.93 (0.70)	4.17 (0.70)	4.21 (0.80)	4.26 (0.84)
2D version of 3D face set	3.49 (0.92)	3.84 (0.83)	4.08 (0.76)	4.23 (0.86)	4.42 (1.01)
Independent 2D face set	3.75 (0.99)	4.00 (0.83)	4.27 (0.82)	4.22 (0.96)	4.32 (1.08)
Muscle					
3D face set	3.47 (0.83)	3.91 (0.63)	4.17 (0.70)	4.27 (0.79)	4.54 (0.88)
2D version of 3D face set	3.10 (0.91)	3.74 (0.80)	4.08 (0.76)	4.42 (0.89)	4.68 (1.06)
Independent 2D face set	3.39 (0.99)	3.66 (0.85)	4.27 (0.82)	4.53(0.92)	4.73(1.12)



$p = 0.665$, $\eta^2 = 0.006$] but a significant interaction was found between transform dimension and transform level [$F(4,264) = 24.75$, $p < 0.001$, $\eta^2 = 0.273$], reflecting a greater impact of muscle transform compared with fat transform on masculinity.

There was a significant interaction between face set and transformed level [$F(8,528) = 2.61$, $p = 0.008$, $\eta^2 = 0.038$]. Further, the three-way interaction among transform dimension, transform level, and face set was significant [$F(8,528) = 2.17$, $p = 0.028$, $\eta^2 = 0.032$]. To understand the three-way interaction, we conducted two-way ANOVA separately for each face set.

3D Face Set

For 3D faces, the main effect of the transform dimension was non-significant [$F(1,66) = 1.36$, $p = 0.252$, $\eta^2 = 0.020$]. There was a significant main effect of transform level [$F(4,264) = 31.17$, $p < 0.001$, $\eta^2 = 0.321$], which was qualified with an interaction between transform dimension and transform level [$F(4,264) = 4.40$, $p = 0.002$, $\eta^2 = 0.062$, see **Figure 4**]. Paired-samples t -tests showed that significant increases in masculinity ratings occurred between all levels of muscle transform ($p \leq 0.004$ each comparison) except between 0 and +2 BMI units ($p = 0.186$). By contrast, there were no significant increases in masculinity ratings for fat transform above normal weight (0, +2.3, and +4 BMI units, $p \geq 0.337$ each comparison). There were significant decreases in masculinity ratings between faces associated with decreased fat mass compared to increased fat mass ($p \leq 0.005$ each comparison). These findings provide further support for our Hypothesis 3 that the facial correlate of muscle mass increases perceived facial masculinity more than the facial correlate of fat mass.

2D Version of 3D Face Set

For the 2D version of the 3D face set, there was no main effect of transform dimension [$F(1,66) = 0.05$, $p = 0.833$, $\eta^2 = 0.001$]. The main effect of transform level was significant [$F(4,264) = 50.85$, $p < 0.001$, $\eta^2 = 0.435$] but was qualified by a significant interaction between transform dimension and transform level [$F(4,264) = 8.63$, $p < 0.001$, $\eta^2 = 0.116$, see **Figure 5**]. Paired-samples t -tests showed an increase in muscle mass by ~2 BMI units significantly increased masculinity ratings throughout the range (−4 to +4 BMI units, $p \leq 0.002$ each comparison). Significant increases in masculinity ratings with fat mass transform were seen in most comparisons ($p \leq 0.014$ each comparison) but no significant increases were seen in comparisons between faces associated with increased fat mass [0 vs. +2.3 BMI units ($p = 0.170$) and +2.3 vs. +4 BMI units ($p = 0.070$)]. These findings are again in line with our prediction that facial correlates of both fat mass and muscle mass positively influence perceived facial masculinity but that also the facial correlate of muscle mass has a larger impact on masculinity.

Independent 2D Face Set

For face transforms based on the independent 2D face set, the main effect of the transform dimension was non-significant [$F(1,66) = 0.02$, $p = 0.888$, $\eta^2 = 0.000$]. A significant main effect of transform level [$F(4,264) = 34.89$, $p < 0.001$, $\eta^2 = 0.346$] reflected faces associated with increased mass (fat or muscle) being considered more masculine.

The interaction between transform dimension and transform level was significant [$F(4,264) = 15.82$, $p < 0.001$, $\eta^2 = 0.193$, see **Figure 6**]. This interaction reflects a greater impact of muscle compared with fat on masculinity ratings. Paired-samples t -tests showed that participants rated faces with higher muscle mass significantly more masculine for comparisons between all five

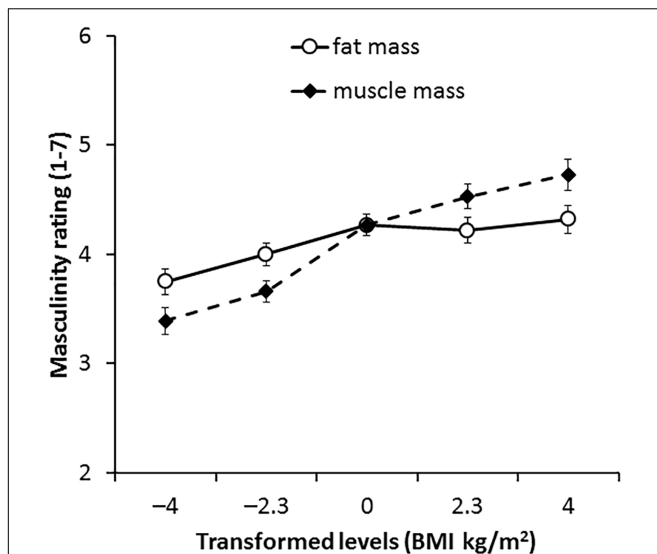


FIGURE 6 | Average masculinity ratings for faces transformed with the face shape correlates of fat and muscle mass for the independent 2D face set. Error bars represent the standard errors.

levels ($p \leq 0.017$ each comparison). In contrast, a significant increase in masculinity ratings for faces associated with higher fat mass was evident only for comparisons between faces with decreased fat mass (-4 BMI units, -2.3 BMI units) and the other levels ($p \leq 0.046$ each comparison). There were no significant differences in masculinity ratings for fat transforms 0 , $+2.3$, or $+4.3$ BMI units ($p \geq 0.270$ each comparison). As fat mass increased from low to normal weight, masculinity increased, but for gain in the fat level above normal weight, there was no significant change in masculinity ratings. These findings support our hypothesis that the facial correlate of muscle mass enhances perceived facial masculinity more than the facial correlate of fat mass.

The interaction between face set, transform dimension and transform level arises from the relative size of the muscle and fat transforms across the three face sets, with the fat and muscle differences being most subtle in the 3D face set though the pattern is similar for each face set.

Discussion

As expected, facial correlates of fat mass and muscle mass both positively affected perceived facial masculinity in men. The results are consistent with Holzleitner et al. (2014) findings of heavier men being perceived as more masculine. As we hypothesized, muscle mass enhances the perception of masculinity more than fat mass. Specifically, increasing the face shape correlate of muscle mass resulted in higher ratings of facial masculinity across the full weight range (BMI range $18-26$). By contrast, increasing the face shape correlate of fat mass only raised masculinity rating from low to normal weight (BMI = $18-22$). Further increases in fat mass above normal weight (BMI = 22) had little or no impact on the perception of masculinity. These results imply that the effect of fat on

masculinity is more prevalent in men with underweight to normal weight bodies.

STUDY 2: ATTRACTION TO THE FACIAL CORRELATES OF BODY COMPOSITION

Study 1 found that facial correlates of both fat mass and muscle mass contribute to perceived facial masculinity, which has been found to affect the perception of attractiveness. In this part of the study, we tested the relationship between facial correlates of body composition and facial attractiveness.

As discussed before, higher levels of masculinity are preferred by women more for short-term relationships than for long-term relationships. Hence, we measured heterosexual women's preferences for facial correlates of body composition in male faces under short-term and long-term relationship contexts. Given the findings above that the facial correlate of muscle mass increases perceived facial masculinity, we predicted that *women would show a stronger preference for the facial correlate of muscle mass in a short-term rather than a long-term relationship context* (Hypothesis 4). Regarding fat mass, in the introduction we hypothesized that women would show a stronger preference for higher fat mass in short-term relationships than in long-term relationships. In the light of the masculinity ratings we found in Study 1, this hypothesis should be modified. We can now hypothesize that *if women show an overall preference for men with a BMI < 22, we predict women will prefer a face shape associated with more fat mass for a short-term relationship in comparison to a long-term relationship* (Hypothesis 5a). Conversely, we predict that *women will not shift their preference for the facial correlate of fat mass between short-term and long-term relationships if they prefer men with a BMI > 22* (Hypothesis 5b). Nevertheless, we predict the preference shift between short-term and long-term contexts will be more apparent for the facial correlate of muscle mass than the facial correlate of fat mass (Hypothesis 6).

This study was initially administered with Study 1 as a single experiment consisting of two tasks (masculinity rating and preference) for University students, with the preference task executed before the masculinity task. Considering the students are highly homogeneous groups due to their age and educational background, the study was repeated in a more heterogeneous group to test the generalizability of findings. Hence, we recruited another group of participants through the online recruitment platform, Amazon MTurk.

Methods

Ethical approval was received from University of St Andrews Ethics Committee (PS13176 and PS13092). Participants gave written informed consent to perform the experimental tasks.

Participants

For the student group, 63 heterosexual female participants ($M_{\text{age}} \pm SD = 18.94 \pm 2.17$, range $18-35$ years; 48 Caucasian) completed this study after exclusion of those without demographic information (age, sex, ethnicity, and sexual orientation) or who reported to be homosexual or males.

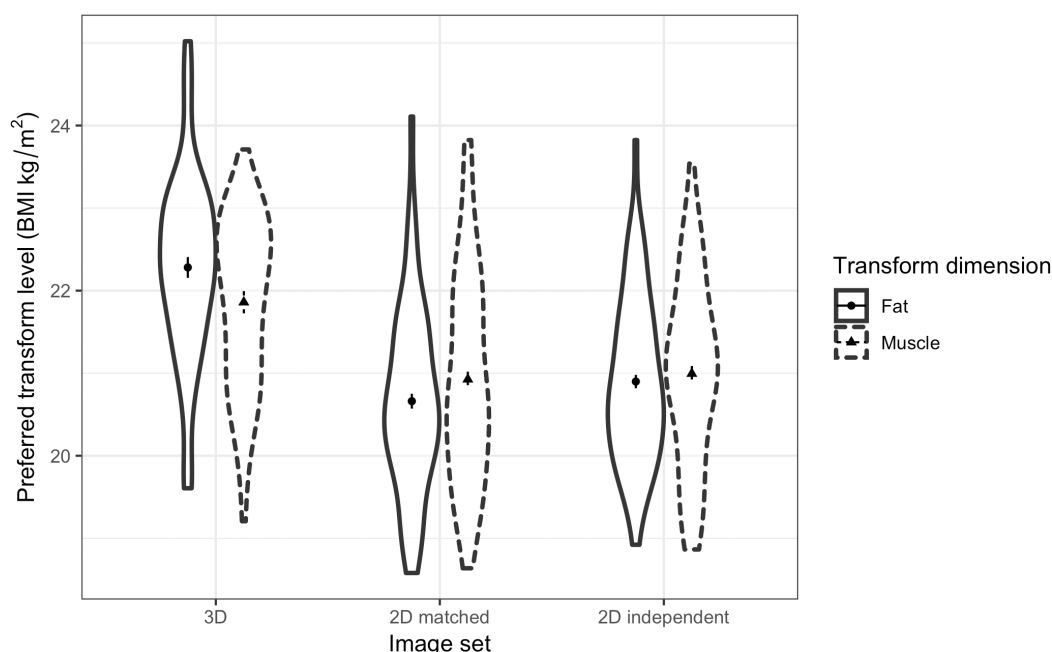


FIGURE 7 | Violin plots showing the distribution of female students' preferences for the facial correlates of fat mass and muscle mass in men. The vertical axis represents the associated BMI of the most preferred faces chosen by the students in short-term and long-term relationship contexts. The error bars represent the standard errors and the symbols indicate means.

For the MTurk workers group, 58 heterosexual women ($M_{\text{age}} \pm SD = 32.09 \pm 6.68$, range 22–45 years; 43 Caucasian) completed this study after exclusion using the same criteria as the students' group and an additional exclusion age criterion. Ten women over age 45 years were additionally excluded as our prediction was based on the assumption that the key benefit women gain from short-term relationships concerns potential reproductive success. MTurk participants were paid \$3 for their time.

Materials

The stimuli consisted of face images transformed as described above. For each face identity, 15 images were produced spanning the transformation ± 4 BMI units on fat mass and muscle mass dimensions. The 15 images were presented as an interactive continuum. For MTurk workers, a total of 30 face continua: 5 face identities \times 2 dimensions (fat/muscle) \times 3 face sets (3D face set, 2D version of 3D face set, independent 2D face set) were presented twice in separate trial blocks asking about preferences for a short-term sexual relationship and long-term relationship. For the student group, the three face identities were used. Thus, 18 face continua: 3 identities \times 2 dimensions (fat/muscle) \times 3 face sets (3D face set, 2D version of 3D face set, independent 2D face set) were presented in each of two trial blocks.

Procedure

At the beginning of this study, participants were asked, "Please indicate the sex of face that you would like to see (as a sexual partner)" (Note: female faces were also given as an option for heterosexual males, homosexual and bisexual female participants

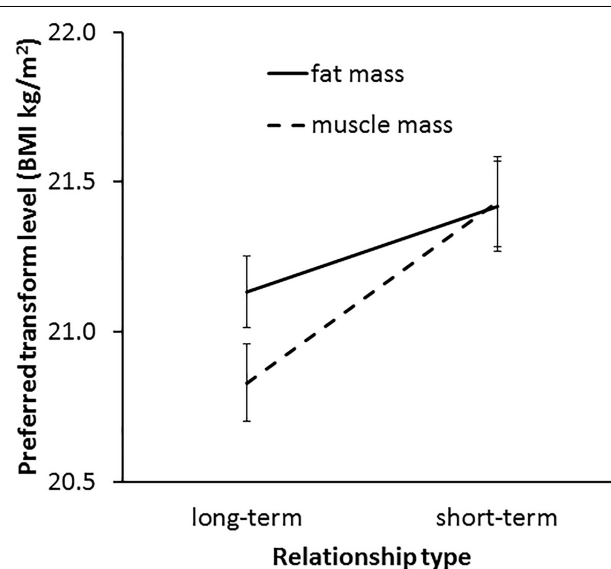


FIGURE 8 | The interaction between relationship type (short-term and long-term) and preferred facial correlates of body composition (fat mass and muscle mass) in student participants. The vertical axis represents the associated BMI of the most preferred faces. Error bars represent standard errors.

to view, but data from these faces are not analyzed here). The participants' demographic information (age, sex, ethnicity, and sexual orientation) was collected in an initial questionnaire. Then

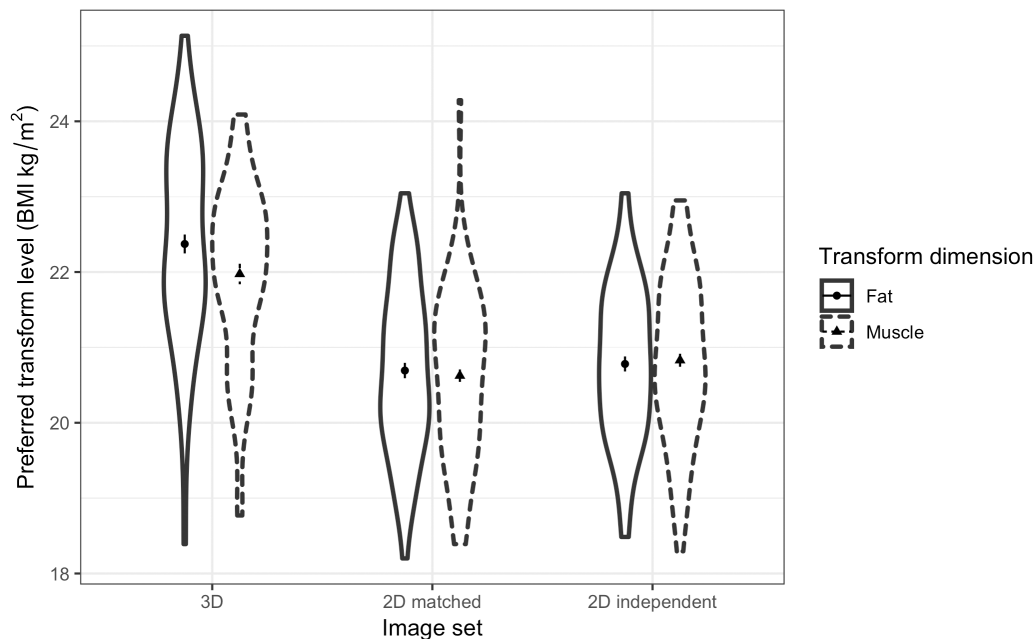


FIGURE 9 | Violin plots showing the distribution of MTurk women's preferences for the facial correlates of fat mass and muscle mass in men. The vertical axis represents the associated BMI of the most preferred faces chosen by the women in short-term and long-term contexts. The error bars represent the standard errors and the symbols indicate means.

participants were presented with the stimuli twice in two blocks. They were asked to adjust the slider underneath each stimulus to make the face most resemble someone they would find attractive as a short-term (sexual) partner and as a long-term partner in two separate blocks. The order of the tasks was counterbalanced. Trials with 2D and 3D face stimuli were also grouped in two separate sub-blocks. The order of sub-blocks and the presentation order within each sub-block was randomized. The scroll direction to change the face shape was randomized across trials. The next image would only be shown when participants adjusted the slider and clicked the submit button. For each trial, the BMI level chosen by each participant was saved.

Instructions were given prior to tasks as follows (a) Short-term (sexual) relationship: "Please change the face to most resemble someone you would find attractive for a SHORT-TERM (sexual) relationship." (b) Long-term relationship: "Please change the face to most resemble someone you would find attractive for a LONG-TERM relationship."

Statistical Analysis

The dependent variable was the transform level that was most preferred (expressed as a BMI equivalent). The data for the students group and MTurk group were analyzed separately in SPSS 24.0.

Results

Student Group

A three-way ANOVA was run to test women's preference for facial correlates of fat mass and muscle mass in different relationship contexts and across the three face sets. The results

showed a non-significant main effect of fat/muscle transform dimension [$F(1,62) = 3.18, p = 0.079, \eta^2 = 0.049$]. As expected, a significant main effect of context [$F(1,62) = 9.26, p = 0.003, \eta^2 = 0.130$] was found, with participants preferring faces of heavier men (with fat mass or muscle mass) for a short-term relationship ($M = 21.42, SD = 1.15$) rather than a long-term relationship ($M = 20.98, SD = 0.90$). In addition, there was a significant main effect of face set [$F(2,124) = 107.37, p < 0.001, \eta^2 = 0.634$, see **Figure 7**]. Although we did not expect to find a main effect of the face set, the paired-samples *t*-tests suggest that the effect might simply be due to participants choosing heavier faces in the 3D face set compared with the other two 2D face sets. Paired-samples *t*-tests showed that participants choose heavier faces for the 3D face set ($M = 22.14, SD = 1.07$) compared with the 2D version of 3D face set ($M = 20.67, SD = 0.99$) [$t(62) = 12.02, p < 0.001$] and the independent 2D face set ($M = 20.80, SD = 0.93$) [$t(62) = 10.88, p < 0.001$].

In line with our Hypothesis 6, a significant interaction was found between transform dimension and context [$F(1,62) = 4.73, p = 0.034, \eta^2 = 0.071$, see **Figure 8**]. This result indicates a greater effect of muscle than fat on preference in the two contexts. As expected, paired-samples *t*-tests showed that a higher level of facial correlate of muscle mass was preferred in a short-term ($M = 21.43, SD = 1.22$) rather than a long-term ($M = 20.83, SD = 1.07$) relationship [$t(62) = 3.49, p = 0.001$]. By contrast, there was a non-significant trend for a difference between preference for the facial correlate of fat mass in short-term ($M = 21.42, SD = 1.23$) and long-term ($M = 21.13, SD = 0.96$) [$t(62) = 1.86, p = 0.068$] relationships, which provides limited support for Hypothesis 5a.

The three-way interaction (transform dimension \times relationship context \times face set) was non-significant [$F(2,124) = 0.33$, $p = 0.719$, $\eta^2 = 0.005$]. Since the interaction between fat and muscle transform and relationship context was found to be significant and independent of the face set, it was not necessary to analyze the data further for each face set separately. Thus, our main prediction was borne out across the three face sets.

Finally, one-sample t -tests compared the preferred BMI (average across the three face sets) with a BMI of 22 (the average of the original starting BMI of the face stimuli) to test whether women show a general preference toward a lower or higher than normal weight. Significant decreases in preferred BMI below 22.0 were found, reflecting a reduction of fat mass and muscle mass for both short-term [fat mass: $M = 21.42$, $t(62) = -3.78$, $p < 0.001$; muscle mass: $M = 21.43$, $t(62) = -3.70$, $p < 0.001$] and long-term [fat mass: $M = 21.13$, $t(62) = -7.18$, $p < 0.001$; muscle mass: $M = 20.83$, $t(62) = -8.72$, $p < 0.001$] relationships.

MTurk Workers

Similarly, a three-way ANOVA was run to test MTurk women's preference for men's facial correlates of fat and muscle mass across relationship contexts. The results showed non-significant main effects of transform dimension [$F(1,57) = 0.06$, $p = 0.808$, $\eta^2 = 0.001$] and context [$F(1,57) = 1.31$, $p = 0.258$, $\eta^2 = 0.022$]. A significant main effect of face set was found [$F(2,114) = 71.58$, $p < 0.001$, $\eta^2 = 0.557$, see Figure 9]. Similar to the student group, paired-samples t -tests showed that participants chose heavier faces (with higher fat mass or muscle mass) with the 3D face set ($M = 22.07$, $SD = 0.89$) compared with the 2D version of 3D face set ($M = 20.79$, $SD = 1.10$) [$t(57) = 8.89$, $p < 0.001$] and the independent 2D face set ($M = 20.95$, $SD = 0.93$) [$t(57) = 8.68$, $p < 0.001$]. Unlike the results from the student group, MTurk participants preferred slightly heavier faces for the independent 2D face set compared to the 2D version of 3D face set [$t(57) = -2.65$, $p = 0.010$].

In line with our Hypothesis 6, a significant interaction was found between fat and muscle transform dimension and relationship context [$F(1,57) = 7.36$, $p = 0.009$, $\eta^2 = 0.114$, see Figure 10]. Paired-samples t -tests results suggest that MTurk women showed a stronger preference for the facial correlate of muscle mass in short-term relationships ($M = 21.42$, $SD = 1.12$) compared with long-term relationships ($M = 21.10$, $SD = 0.95$) [$t(57) = 2.33$, $p = 0.024$] but those women did not differ in their preference for the facial correlate of fat mass between short-term ($M = 21.24$, $SD = 0.99$) and long-term relationships ($M = 21.32$, $SD = 0.96$) [$t(57) = 0.70$, $p = 0.488$]. Further, the three-way interaction (transform dimension \times relationship context \times face set) was non-significant [$F(2,114) = 1.52$, $p = 0.224$, $\eta^2 = 0.026$], indicating that the interaction between fat/muscle transform and relationship context was consistent across the three face sets.

One-sample t -tests compared the preferred BMI transform level (average across the three face sets) to a BMI of 22 (the average of the original starting BMI of the face stimuli). MTurk participants preferred a BMI significantly reduced from a BMI of 22 for both fat mass and muscle mass in short-term [fat mass: $M = 21.24$, $t(57) = -5.82$, $p < 0.001$; muscle mass: $M = 21.42$,

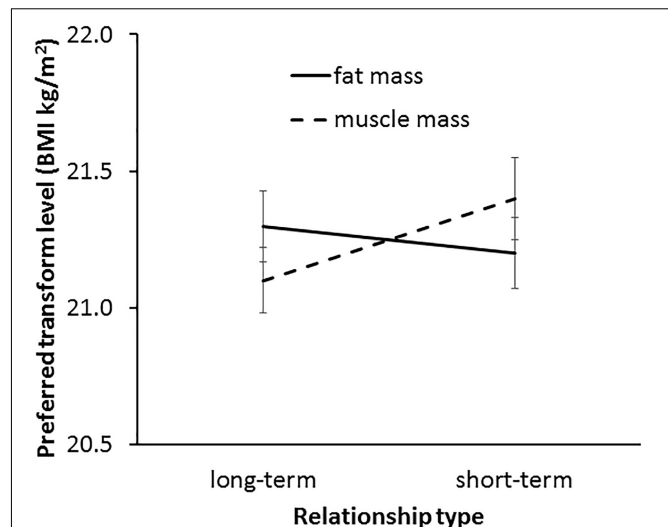


FIGURE 10 | The interaction between relationship context (short-term vs. long-term relationship) and preferred facial correlates of body composition (fat mass and muscle mass) in MTurk participants. The vertical axis represents the associated BMI of the most preferred faces. Error bars represent standard errors.

$t(57) = -3.96$, $p < 0.001$] and long-term [fat mass: $M = 21.32$, $t(57) = -5.37$, $p < 0.001$; muscle mass: $M = 21.10$, $t(57) = -7.17$, $p < 0.001$] relationships.

Discussion

This study investigated heterosexual women's preferences for men's facial correlates of body composition under different relationship contexts. In line with our Hypothesis 4, women showed a stronger preference for faces associated with higher muscle mass in a short-term relationship compared with a long-term relationship. In contrast, women did not shift their preference for the facial correlate of fat mass between short-term and long-term relationships even though their overall preference lay in the low end of normal weight (BMI ~21 kg/m²).

GENERAL DISCUSSION

The present study had two aims: first, to investigate the effect of facial correlates of body composition (fat mass and muscle mass) on the perceived facial masculinity of men and second to investigate the effect of facial correlates of body composition on women's preferences in different relationship contexts. Ratings of masculinity supported our hypotheses that both facial correlates of fat mass and muscle mass positively affect perceived facial masculinity. While the facial correlate of muscle mass had a pronounced effect on perceived masculinity, the effect of the facial correlate of fat mass increased masculinity only in underweight to lower normal weight men. In interactive preferences tests where women optimized the shape of a male face, we found that there is a context shift in preferences with women preferring facial correlates of higher muscle mass for a short-term relationship compared to a long-term relationship. By

contrast, we found that women do not shift their preference for the facial correlate of fat mass between short-term and long-term relationships.

Attribution to Perceived Facial Masculinity

The results from Study 1 supported our predictions that facial correlates of body composition influence perceived facial masculinity. In line with Holzleitner et al. (2014) findings, the facial cues to higher body weight (BMI) increase perceived facial masculinity of male faces. The results extend previous findings that 'facial adiposity' (weight perceived from the facial appearance) is positively associated with perceived masculinity in under to normal weight men but not in overweight or obese men (Phalane et al., 2017). It should be noted that the definition of facial adiposity in Phalane et al.'s (2017) study was a measure of the weight perceived from the face. Hence, their perceived adiposity measure will include two components, namely weight from fat and weight from muscle. Phalane et al.'s (2017) results indicate a quadratic relationship between perceived facial adiposity and masculinity. By distinguishing the facial correlates of fat and muscle, we find a quadratic relationship between fat and masculinity, but a linear relationship between muscle and masculinity. Hence our study shows that the findings of Phalane et al. (2017) are likely to reflect the facial correlate of fat. Our findings indicate that the muscle and fat components should be treated separately in future work on facial perception.

Importantly, our results were consistent across the three face sets employed. Although the relationship between the facial correlate of fat mass and masculinity was slightly different between the 2D version of 3D face set and the other two face sets, the facial correlate of muscle mass was found to have a larger impact on perceived masculinity across all three sets of faces.

The distinct effects of fat mass and muscle mass on perceived facial masculinity might reflect the sex differences in body physique because men are generally heavier in body weight and have more muscle mass than women (Wells, 2007). Indeed, fat-free muscle mass are even more sexually dimorphic than differences in body weight (Lassek and Gaulin, 2009). Hence, heavier men with higher muscle mass have attributes associated with higher sexual dimorphism and should be seen as more masculine. Indeed, this is what we found in the first part of our study. Although men on average have greater weight compared to women, the weight difference is mainly due to the higher muscle mass that men possess. Hence, the excess fat mass does not make male faces more masculine but decreased weight, whether due to loss of fat mass or loss of muscle mass, decreases men's perceived masculinity.

It is also possible that the facial correlates of muscle serve as a cue to testosterone levels and thus enhance masculinity perception more than the facial correlate of fat mass. In fact, increased testosterone levels during puberty cause growth of jaw, brow, chin and nose (Marečková et al., 2011). As a result, adult male faces have a relatively longer and broader lower jaw, higher brow ridges, thinner cheeks and more prominent cheekbones compared to adult women (Little et al., 2011b). The

perceptual studies here provide further evidence that the face shape correlates of fat mass and muscle mass are distinct in men. Holzleitner and Perrett (2016) found that observers were able to distinguish the face shape correlates of fat mass and muscle mass using 3D facial stimuli. Here, we find further distinctions for the fat and muscle aspects of body composition for both 2D and 3D facial stimuli. A visual adaptation study also suggested that body fat and muscle are processed independently in the brain (Sturman et al., 2017). The face shape correlates of muscle may not only provide cues to body composition and physique but also may provide a cue to testosterone levels, and hence influence masculinity perception.

Taken together, we have shown that the perception of male facial masculinity is not only based on the cues to body weight. More importantly, muscularity is the aspect of the body composition that has greatest influence on facial masculinity perception.

Context Shifts in Preferences for Facial Masculinity

Study 2 indicates that women's preference for male face shape is dependent on context: we found that women preferred faces associated with a higher muscle mass for short-term relationships rather than long-term relationships but that women do not show different preferences for facial cues to fat mass between short- and long-term relationships.

Our findings appear to be in line with the good genes hypothesis, which argues that women are attracted to indicators signaling heritable aspects of immunity and health when seeking short-term partners (Gangestad and Simpson, 2000; Gangestad et al., 2005). We note that the contextual differences in preferences are also consistent with an alternative interpretation that the preference difference might reflect avoidance of negative characteristics associated with higher muscularity in long-term relationships. Previous studies have revealed that men with high testosterone levels and more fat-free mass (greater muscle mass) report having a larger number of sex partners, indicating that these men might devote more effort in mating relative to parenting (Peters et al., 2008; Lassek and Gaulin, 2009). Further, other studies show that men with high testosterone levels are less likely to get married and more likely to get divorced (Julian and McKenry, 1989; Booth and Dabbs, 1993; Booth et al., 2000). Hence, male faces that reflect high levels of androgen-mediated traits may be less preferred by women in a long-term relationship because of the associated behavioral traits that are inconsistent with paternal investment.

This interpretation may also account for why women do not show different preferences for the facial correlate of fat mass between the two relationship contexts. Although we predicted facial cues to higher fat mass would be preferred for short-term relationships because higher fat mass contributes to facial masculinity (at least in low weight men), the masculinity perception contributed by the facial correlate of fat mass, however, is not testosterone dependent. Therefore, despite the fact that faces associated with higher fat mass are perceived to be more masculine, the same facial cues to fat mass are

not necessarily associated with the undesirable testosterone-mediated traits. Consequently, women do not need to shift their preference between short- and long-term relationships since there are no (or fewer) associated costs with preferring masculinity that derives from slightly higher fat mass. Therefore, the relationship context preference differences that we find may reflect women's reluctance to choose very muscular men who appear unsuitable as long-term partners. Future studies investigating the perception of personality traits from facial cues to fat mass and muscle mass may provide better understandings for the context shifts.

It worth mentioning that women generally prefer faces reflecting low fat mass and muscle mass under both contexts. The associated BMI of the most preferred face was significantly reduced compared with the original starting BMI of the facial stimuli (namely BMI of 22.0 kg/m²). This suggests that men with low-normal body weight but not underweight are most preferred by women as partners. This finding is in line with previous studies on men's attractiveness and BMI, which found that the most preferred male bodies resemble BMI around 21 kg/m² (Swami and Tovée, 2005, 2008). The findings are also consistent with one prior study, which found an inverted U shape relationship between men's body attractiveness and muscularity (Frederick and Haselton, 2007). Men with medium levels of muscle mass were rated to be more sexually desirable compared with the very low or very high levels of muscularity (Frederick and Haselton, 2007).

By contrast, our findings are less consistent with recent findings that stronger men are seen as more attractive (Sell et al., 2017; Foo et al., 2018) with a linear increase in attractiveness reported for the range of men's strength sampled. There are two possible reasons for the inconsistency. Firstly, it should be noted that the studies mainly focused on attractiveness of men's bodies rather than men's faces. There might be a discrepancy between the attractiveness of men's bodies and faces. Women might find a stronger body attractive but not necessarily the face shape accompanying such a body. Future study may set out to test whether women show consistent preferences for men's body muscularity and the facial correlates of muscle.

Second, the studies that found a positive relationship between strength and attractiveness have adopted a correlational method comparing strength to ratings of natural bodies (Sell et al., 2017; Foo et al., 2018), while we employed an interactive method to let participants optimize the most attractive face shape from stimuli synthesized with computer graphics. Support for the divergence of results reflecting different methods comes from the study of Brierley et al. (2016) who used a similar interactive method to test the attractiveness of men's bodies. Brierley et al. (2016) found that a slight decrease of body fat and slight increase of body muscle was optimal for men with normal starting BMI and body composition. In both the experiment of Brierley et al. (2016) and the experiment here, men with a high muscular body composition were not the most attractive. Studies comparing ratings of real and computer-manipulated images may help resolve the difference in attraction of strong and muscular men.

Although our hypotheses are supported with the use of both 2D and 3D facial stimuli, we note that a higher BMI (in both fat and muscle dimensions) was preferred in 3D faces compared to 2D faces. This effect of dimensionality might be due to the fact that our 3D stimuli combined both the front and the profile views, whereas our 2D stimuli used the front view alone. The combination of front and profile views may provide more information relating to weight. Alternatively, the profile view may provide information that is distinct from that evident in the front view. Indeed, prior study has shown that women make different choices for attractiveness and dominance when viewing front and profile views of the male faces (Swaddle and Reiersen, 2002). Furthermore, Danel et al. (2018) showed that the measured sexually dimorphic facial features show only a moderate correlation across front and profile views ($r = 0.20$). These findings imply that further experiments are required to understand the processing of frontal and lateral views of the face.

CONCLUSION

In summary, we have shown the distinct effects that facial correlates of fat mass and muscle mass have on perceptions of masculinity and attractiveness in men. Our findings show that the facial correlate of muscle mass has a profound impact on perceived facial masculinity in men of all weights. By contrast, the facial correlate of fat mass affects masculinity only in underweight to lower normal weight men. Further, we find a contextual shift in women's attraction to the facial correlate of muscle mass but not fat mass, with a stronger preference for male face shapes associated with high muscle mass under a short-term relationship context compared to a long-term relationship context.

Body size has an impact on a variety of social judgments including attractiveness, strength, dominance, leadership and employment (Windhager et al., 2011; Re and Perrett, 2014; Holzleitner and Perrett, 2016; Nickson et al., 2016; Phalane et al., 2017). Our findings highlight the importance of differentiating size-related effects separately for body fat and body muscle.

In spite of consistent results across the three face sets and two samples of participants, we note that the current studies used a limited number of face identities that were restricted to Caucasian ethnicity. A large and more diverse sample of faces should be employed in future studies.

AUTHOR CONTRIBUTIONS

XL and DP conceived and designed the studies, analyzed the data, and wrote the manuscript. XL contributed to 2D stimuli production and data collection. IH and XL contributed to 3D stimuli production.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02658/full#supplementary-material>

REFERENCES

- Batres, C., and Perrett, D. I. (2017). How the harsh environment of an army training camp changes human (*Homo sapiens*) facial preferences. *Ethology* 123, 61–68. doi: 10.1111/eth.12571
- Booth, A., Carver, K., and Granger, D. A. (2000). Biosocial perspectives on the family. *J. Marriage Fam.* 62, 1018–1034. doi: 10.1111/j.1741-3737.2000.01018.x
- Booth, A., and Dabbs, J. M. Jr. (1993). Testosterone and men's marriages. *Soc. Forces* 72, 463–477. doi: 10.1093/sf/72.2.463
- Boothroyd, L. G., Jones, B. C., Burt, D. M., and Perrett, D. I. (2007). Partner characteristics associated with masculinity, health and maturity in male faces. *Pers. Individ. Dif.* 43, 1161–1173. doi: 10.1016/j.paid.2007.03.008
- Brierley, M. E., Brooks, K. R., Mond, J., Stevenson, R. J., and Stephen, I. D. (2016). The body and the beautiful: health, attractiveness and body composition in men's and women's bodies. *PLoS One* 11:e0156722. doi: 10.1371/journal.pone.0156722
- Buss, D. M., and Schmitt, D. P. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychol. Rev.* 100, 204–232. doi: 10.1037/0033-295X.100.2.204
- Confer, J. C., Perilloux, C., and Buss, D. M. (2010). More than just a pretty face: men's priority shifts toward bodily attractiveness in short-term versus long-term mating contexts. *Evol. Hum. Behav.* 31, 348–353. doi: 10.1016/j.evolhumbehav.2010.04.002
- Cornwell, R. E., Boothroyd, L. G., Burt, D. M., Feinburg, D. R., Jones, B. C., Little, A. C., et al. (2004). Concordant preferences for opposite-sex signals? Human pheromones and facial characteristics. *Proc. R. Soc. Lond. B Biol. Sci.* 271, 635–640. doi: 10.1098/rspb.2003.2649
- Currie, T. E., and Little, A. C. (2009). The relative importance of the face and body in judgments of human physical attractiveness. *Evol. Hum. Behav.* 30, 409–416. doi: 10.1016/j.evolhumbehav.2009.06.005
- Danel, D. P., Valentova, J. V., Sanchez, O. R., Leongómez, J. D., Varella, M. A. C., and Kleisner, K. (2018). A cross-cultural study of sex-typicality and averageness: correlation between frontal and lateral measures of human faces. *Am. J. Hum. Biol.* 30:e23147. doi: 10.1002/ajhb.23147
- DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proc. R. Soc. Lond. B Biol. Sci.* 273, 1355–1360. doi: 10.1098/rspb.2005.3445
- DeBruine, L. M., Jones, B. C., Smith, F. G., and Little, A. C. (2010). Are attractive men's faces masculine or feminine? The importance of controlling confounds in face stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 751–758. doi: 10.1037/a0016457
- Elagizi, A., Köhler, T. S., and Lavie, C. J. (2018). Testosterone and cardiovascular health. *Mayo Clin. Proc.* 93, 83–100. doi: 10.1016/j.mayocp.2017.11.006
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., and Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. *Evol. Hum. Behav.* 29, 233–241. doi: 10.1016/j.evolhumbehav.2007.12.008
- Folstad, I., and Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *Am. Nat.* 139, 603–622. doi: 10.1086/285346
- Foo, Y. Z., Nakagawa, S., Rhodes, G., and Simmons, L. W. (2017). The effects of sex hormones on immune function: a meta-analysis. *Biol. Rev.* 92, 551–571. doi: 10.1111/brv.12243
- Foo, Y. Z., Simmons, L. W., Marianne, P., and Rhodes, G. (2018). Perceived physical strength in men is attractive to women but may come at a cost to ejaculate quality. *Anim. Behav.* 142, 191–197. doi: 10.1016/j.anbehav.2018.06.019
- Frederick, D. A., and Haselton, M. G. (2007). Why is muscularity sexy? Tests of the fitness indicator hypothesis. *Pers. Soc. Psychol. Bull.* 33, 1167–1183. doi: 10.1177/0146167207303022
- Gangestad, S. W., and Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *Behav. Brain Sci.* 23, 573–587. doi: 10.1017/S0140525X0000337X
- Gangestad, S. W., Thornhill, R., and Garver-Apgar, C. E. (2005). Adaptations to ovulation: implications for sexual and social behavior. *Curr. Dir. Psychol. Sci.* 14, 312–316. doi: 10.1111/j.0963-7214.2005.00388.x
- Holzleitner, I. J., Hunter, D. W., Tiddeman, B. P., Seck, A., Re, D. E., and Perrett, D. I. (2014). Men's facial masculinity: when (body) size matters. *Perception* 43, 1191–1202. doi: 10.1068/p7673
- Holzleitner, I. J., and Perrett, D. I. (2016). Perception of strength from 3D faces is linked to facial cues of physique. *Evol. Hum. Behav.* 37, 217–229. doi: 10.1016/j.evolhumbehav.2015.11.004
- Holzleitner, I. J., and Perrett, D. I. (2017). Women's preferences for men's facial masculinity: trade-off accounts revisited. *Adapt. Hum. Behav. Physiol.* 3, 304–320. doi: 10.1007/s40750-017-0070-3
- Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., et al. (2018). No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. *Psychol. Sci.* 29, 996–1005. doi: 10.1101/155788
- Julian, T., and McKenry, P. C. (1989). Relationship of testosterone to men's family functioning at mid-life: a research note. *Aggress. Behav.* 15, 281–289. doi: 10.1002/ab.2480150403
- Kelly, D. M., and Jones, T. H. (2013). Testosterone: a metabolic hormone in health and disease. *J. Endocrinol.* 217, 25–45. doi: 10.1530/JOE-12-0455
- Lassek, W. D., and Gaulin, S. J. (2009). Costs and benefits of fat-free muscle mass in men: relationship to mating success, dietary requirements, and native immunity. *Evol. Hum. Behav.* 30, 322–328. doi: 10.1016/j.evolhumbehav.2009.04.002
- Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C., and Roberts, S. C. (2011a). Human preference for masculinity differs according to context in faces, bodies, voices, and smell. *Behav. Ecol.* 22, 862–868. doi: 10.1093/beheco/arr061
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011b). Facial attractiveness: evolutionary based research. *Phil. Trans. R. Soc. B Bio. Sci.* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Little, A. C., Jones, B. C., DeBruine, L. M., and Feinberg, D. R. (2008). Symmetry and sexual-dimorphism in human faces: interrelated preferences suggest both signal quality. *Behav. Ecol.* 19, 902–908. doi: 10.1093/beheco/arn049
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. R. Soc. Lond. B Biol. Sci.* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Marečková, K., Weinbrand, Z., Chakravarty, M. M., Lawrence, C., Aleong, R., Leonard, G., et al. (2011). Testosterone-mediated sex differences in the face shape during adolescence: subjective impressions and objective features. *Horm. Behav.* 60, 681–690. doi: 10.1016/j.yhbeh.2011.09.004
- Mooradian, A. D., Morley, J. E., and Korenman, S. G. (1987). Biological actions of androgens. *Endoc. Rev.* 8, 1–28. doi: 10.1210/edrv-8-1-1
- Nickson, D., Timming, A. R., Re, D., and Perrett, D. I. (2016). Subtle increases in BMI within a healthy weight range still reduce women's employment chances in the service sector. *PLoS One* 11:e0159659. doi: 10.1371/journal.pone.0159659
- Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P., and Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*). *J. Comp. Psychol.* 117, 264–271. doi: 10.1037/0735-7036.117.3.264
- Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., et al. (1999). Menstrual cycle alters face preference. *Nature* 399, 741–742. doi: 10.1038/21557
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Peters, M., Simmons, L. W., and Rhodes, G. (2008). Testosterone is associated with mating success but not attractiveness or masculinity in human males. *Anim. Behav.* 76, 297–303. doi: 10.1016/j.anbehav.2008.02.008
- Phalane, K. G., Tribe, C., Steel, H. C., Cholo, M. C., and Coetzee, V. (2017). Facial appearance reveals immunity in African men. *Sci. Rep.* 7:7443. doi: 10.1038/s41598-017-08015-9
- Re, D. E., and Perrett, D. I. (2014). The effects of facial adiposity on attractiveness and perceived leadership ability. *Q. J. Exp. Psychol.* 67, 676–686. doi: 10.1080/17470218.2013.825635
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208

- Rhodes, G., Chan, J., Zebrowitz, L. A., and Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proc. R. Soc. Lond. B Biol. Sci.* 270, 93–95. doi: 10.1348/000712600161718
- Saxton, T. K., Little, A. C., Rowland, H. M., Gao, T., and Roberts, S. C. (2009). Trade-offs between markers of absolute and relative quality in human facial preferences. *Behav. Ecol.* 20, 1133–1137. doi: 10.1093/beheco/arp107
- Scott, I. M. L., Pound, N., Stephen, I. D., Clark, A. P., and Penton-Voak, I. S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PLoS One* 5:e13585. doi: 10.1371/journal.pone.0013585
- Sell, A., Lukaszewski, A. W., and Townsley, M. (2017). Cues of upper body strength account for most of the variance in men's bodily attractiveness. *Proc. R. Soc. B* 284:20171819. doi: 10.1098/rspb.2017.1819
- Sturman, D., Stephen, I. D., Mond, J., Stevenson, R. J., and Brooks, K. R. (2017). Independent aftereffects of fat and muscle: implications for neural encoding, body space representation, and body image disturbance. *Sci. Rep.* 7:40392. doi: 10.1038/srep40392
- Swaddle, J. P., and Reiersen, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proc. R. Soc. Lond. B Biol. Sci.* 269, 2285–2289. doi: 10.1098/rspb.2002.2165
- Swami, V., and Tovée, M. J. (2005). Male physical attractiveness in Britain and Malaysia: a cross-cultural study. *Body Image* 2, 383–393. doi: 10.1016/j.bodyim.2005.08.001
- Swami, V., and Tovée, M. J. (2008). The muscular male: a comparison of the physical attractiveness preferences of gay and heterosexual men. *Int. J. Mens Health* 7, 59–71. doi: 10.3149/jmh.0701.59
- Tiddeman, B., Burt, M., and Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Comput. Graph. Appl.* 21, 42–50. doi: 10.1109/38.946630
- Wells, J. C. (2007). Sexual dimorphism of body composition. *Best Pract. Res. Clin. Endocrinol. Metab.* 21, 415–430. doi: 10.1016/j.beem.2007.04.007
- Windhager, S., Schaefer, K., and Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *Am. J. Hum. Biol.* 23, 805–814. doi: 10.1002/ajhb.2121

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Lei, Holzleitner and Perrett. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Predictors of Fighting Ability Inferences Based on Faces

Vít Třebický^{1,2*}, Jitka Fialová^{1,2†}, David Stella^{1,2}, Klára Coufalová³, Radim Pavelka³, Karel Kleisner^{1,2}, Radim Kuba¹, Zuzana Štěrbová^{1,2} and Jan Havlíček^{1,2}

¹ Faculty of Science, Charles University, Prague, Czechia, ² Applied Neurosciences and Brain Imaging, National Institute of Mental Health, Klecany, Czechia, ³ Faculty of Physical Education and Sport, Charles University, Prague, Czechia

OPEN ACCESS

Edited by:

Danielle Sulikowski,
Charles Sturt University, Australia

Reviewed by:

Justin Kyle Mogilski,
University of South Carolina
Salkehatchie, United States
Barnaby James Wyld Dixon,
The University of Queensland,
Australia

*Correspondence:

Vít Třebický
vit.trebicky@natur.cuni.cz

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 October 2018

Accepted: 19 December 2018

Published: 15 January 2019

Citation:

Třebický V, Fialová J, Stella D,
Coufalová K, Pavelka R, Kleisner K,
Kuba R, Štěrbová Z and Havlíček J
(2019) Predictors of Fighting Ability
Inferences Based on Faces.
Front. Psychol. 9:2740.
doi: 10.3389/fpsyg.2018.02740

Facial perception plays a key role in various social interactions, including formidability assessments. People make relatively accurate inferences about men's physical strength, aggressiveness, and success in physical confrontations based on facial cues. The physical factors related to the perception of fighting ability and their relative contribution have not been investigated yet, since most existing studies employed only a limited number of threat potential measures or proxies. In the present study, we collected data from Czech Mixed Martial Arts (MMA) fighters regarding their fighting success and physical performance in order to test physical predictors of perceived fighting ability made on the basis of high-fidelity facial photographs. We have also explored the relationship between perceived and actual fighting ability. We created standardized 360° photographs of 44 MMA fighters which were assessed on their perceived fighting ability by 94 raters (46 males). Further, we obtained data regarding their physical characteristics (e.g., age, height, body composition) and performance (MMA score, isometric strength, anaerobic performance, lung capacity). In contrast to previous studies, we did not find any significant links between the actual and the perceived fighting ability. The results of a multiple regression analysis have, however, shown that heavier fighters and those with higher anaerobic performance were judged as more successful. Our results suggest that certain physical performance-related characteristics are mirrored in individuals' faces but assessments of fighting success based on facial cues are not congruent with actual fighting performance.

Keywords: perception, formidability, aggressiveness, strength, anaerobic performance, vital capacity, body composition, beardedness

INTRODUCTION

Male intra-sexual competition is considered an important factor of selective pressure (Puts, 2010; Třebický et al., 2012; Hill et al., 2013; Sell et al., 2017), because it is associated with access to resources via rise in social hierarchy and consequently also with broader mating opportunities. Evidence from various cultures (e.g., von Rueden et al., 2008) and ancestral societies (Walker, 2001) suggests that incidence of physical confrontations in humans is comparable to non-human species (Ellis, 1995). Benefits that can be gained in such confrontations must be, however, always weighed against potential costs, which may include injuries or even death. Decision whether to flee or fight is therefore frequently taken before an actual physical confrontation takes place, which means that one of the opponents often surrenders

without a fight (Sell et al., 2012). Individuals who are good at assessing their chances to win are likely to gain a selective advantage. We may thus expect that perceptual and/or cognitive adaptations for the assessments of one's own and others' fighting ability have evolved.

Recent research shows that humans are capable of inferring fighting ability from facial, body, and vocal cues (Sell et al., 2009, 2010; Puts et al., 2012; Třebický et al., 2013; Little et al., 2015; Raine et al., 2018). Current studies tend to focus on investigating the relationship between the individual components of threat potential, such as body size, upper-body strength, or fighting success, and facial perception (Sell, 2016). One cross-cultural study demonstrated that people can assess upper-body strength and fighting ability of males from facial photographs alone (Sell et al., 2009). Several other studies have investigated the association between hand-grip strength—a frequently used proxy for upper-body strength (Wind et al., 2010)—and various characteristics perceived from faces. It has been repeatedly shown that physically stronger men receive higher ratings of dominance, masculinity, and attractiveness (Fink et al., 2007; Windhager et al., 2011; Geniole and McCormick, 2015; Gallup and Fink, 2018). When 3D facial stimuli were used, Holzleitner and Perrett (2016) found an association between actual and perceived strength, but weaker than in earlier investigations. Results of this study also suggest that perceived strength was independently predicted by the amount of muscle and fat, which mediated the effect of actual strength on the perceived strength (Holzleitner and Perrett, 2016). A recent study revealed that men's perceived "facial threat potential"—derived from dominance, strength, and weight ratings—is related to scores of "actual threat potential," as based on a composite measure of hand-grip strength, weight, and height (Han et al., 2017).

Another line of research investigates the association between actual fighting ability and facial perception by employing Mixed Martial Arts (MMA) fighters and their fighting success score. When fight outcomes were assessed from faces of particular pairs of MMA fighters with known fight outcome, the actual winners were selected as more likely to win a fight, as being more aggressive, stronger, and more masculine than the losers (Little et al., 2015). A rating study by Třebický et al. (2013) showed that perceived aggressiveness of MMA fighters is associated with their fighting success. Moreover, actual fighting ability was also linked to perceived fighting ability, but only in heavyweight fighters. However, the factors responsible for the perception of fighting ability, including their relative contribution, has not been investigated yet. This is partly due to the fact that most existing studies employ only a limited number of threat potential measures and/or proxies.

To explore these issues, we collected detailed data on Czech MMA fighters, regarding their actual fighting ability, and those physical performance measurements which were considered important in previous studies focused on the performance of MMA fighters (e.g., Lenetsky and Harris, 2012; Alm and Yu, 2013). We have chosen MMA as an analog to real-life physical confrontations because it combines various fighting styles used in other combat sports and blends them into a unique multielement martial art. It employs a wide variety of techniques: opponents

fight in a standing position, where they rely on punches and kicks (much like in boxing, kick-boxing, and Muay Thai), but also on the ground, where they wrestle and grapple (using techniques from e.g., Brazilian Jiu-Jitsu, Judo, Greco-Roman wrestling, and freestyle wrestling). The extremely dynamic nature of MMA fights involves both repeated explosive movements and submaximal dynamic work, that is, a combination of high anaerobic and aerobic demands (Lenetsky and Harris, 2012). For these reasons, body composition (Boileau and Lohman, 1977; Braswell et al., 2010), aerobic endurance (Yoon, 2002; Radovanovic et al., 2011; Durmic et al., 2017), maximum strength, and anaerobic capacity (AC) (La Bountry et al., 2011) all play an important role in maintaining performance throughout the fight.

To cover a broad range of physical factors which might affect perceived fighting ability, we collected data on overall body strength (measured as the maximal isometric strength of hands, arms, legs, trunk, and neck), endurance (using lung capacity measurements), AC (using the Wingate test), and body composition (data on body weight, body fat mass, muscle mass, and bone mass).

Further, it has been shown that men's beardedness, while having no effect on fighting outcomes in competitions (Dixson et al., 2018), is linked to judgements of higher levels of masculinity (Dixson et al., 2017), dominance (Muscarella and Cunningham, 1996; Neave and Shields, 2008; Dixson and Vasey, 2012; Saxton et al., 2016; Sherlock et al., 2017), and aggressiveness (Muscarella and Cunningham, 1996; Neave and Shields, 2008; Dixson and Vasey, 2012; Geniole and McCormick, 2015). For this reason, we have also explored the effect of facial hair on the perception of fighting ability.

Most existing studies tended to rely on static frontal facial photographs of varying quality and standardization, which convey a limited amount of visual information regarding overall facial morphology (Danel et al., 2018). To overcome these issues, we collected highly standardized 360° view photographs of heads, which provide more visual information. These were then used to investigate the relationship between the perception of fighting ability and various measures of athletes' physical performance.

MATERIALS AND METHODS

All procedures followed were in accordance with ethical standards of the relevant committee on human experimentation and with the Helsinki Declaration. The study was approved by the Institutional Review Board of the National Institute of Mental Health, Czech Republic (Ref. num. 28/15). All participants were informed about the goals of the study and approved their participation by signing informed consent.

Participants

Targets

In total, we obtained photographs and data on physical performance from 44 MMA athletes (mean age = 26.7, $SD = 5.91$, range = 18–38); all residents of the Czech Republic. They were recruited via social media advertisements, leaflets distributed at local MMA tournaments, at gyms, and with

the help of Mixed Martial Arts Association Czech Republic (MMAA). They were reimbursed for their participation with 400 CZK (~15 EUR). To obtain information about their fighting success rate, we computed their actual fighting ability as the proportion of wins relative to the total number of fights. Hereafter, the term “actual fighting ability” refers to their actual success in competition.

Raters

Photographs were judged by 46 male (mean age = 21.96 years, $SD = 2.56$, range = 19–29) and 48 female raters (mean age = 22.29 years, $SD = 3.56$, range = 18–38), mainly Charles University students, who were recruited via social media advertisements and mailing list of participants established in previous studies. The participants received 100 CZK (~4 EUR) as a compensation for their participation and a debriefing leaflet which explained the purpose of the study.

Stimuli Collection

Photographs Acquisition and Setting

Photographs were captured with 24 megapixels full-frame (35.9×24 mm CMOS sensor, a 35 mm film equivalent) digital SLR camera Nikon D610 equipped with a fixed focal length lens Nikon AF-S NIKKOR 85 mm f/1.8 G (Třebický et al., 2016). Exposure values were set to ISO 100, shutter speed 1/200 s and aperture f11 in all photographs. Photographs were shot into 14-bit uncompressed raw files (NEF) and AdobeRGB color space. Color calibration was performed by using X-Rite Color Checker Passport color targets and a white balance patch photographed at the beginning of each session. The camera was mounted in portrait orientation directly on the light stand which also carried a strobe light positioned 125 cm from the participant. The aim of this setup was to achieve a perception close to the social interpersonal distance (Hall, 1966; Baldassare and Feller, 1975; Sorokowska et al., 2017), to maintain a constant distortion of perspective (Třebický et al., 2016; Erkelens, 2018), and to avoid potential perception bias based on interpersonal distance (Bryan et al., 2012). Camera's distance from each participant was checked with a digital laser rangefinder (Bosch PLR 15). Camera's height was adjusted individually for each participant (target) so as to center his head in the middle of the frame, and the focus point was set on participant's eye. This setting of camera's distance, focal length, and sensor size yielded a 35×53 cm field of view (23.85° angle of view).

Participants were seated on a bar stool and asked to sit straight with hands hanging freely alongside their body. They were photographed in black underwear shorts we provided them with (i.e., they wore no T-shirts) and without any adornments, such as glasses or jewelry. They were instructed to look directly into the camera, adopt a neutral expression, and retain this position through whole photographing session. To capture 360° images of targets' head, a stool was placed on a turning platform which could be manually rotated by 10° in 36 steps, see **Figure 1** for illustration. This resulted in 36 photographs for each participant. The platform was placed in a purpose-built portable photographic booth to control for any changes in ambient light and for color reflections (Rowland and Burriss,



FIGURE 1 | Illustrative image of photograph acquisition setup. Photograph by Jitka Fialová, published with informed and written consent of depicted participant and co-authors.

2017; Thorstenson, 2018). We took two full rotations of each participant to obtain one full set of high-quality photographs (to eliminate possible movements between shots, blinks etc.).

Standardized lighting conditions and uniform exposure across the whole scene were ensured by using one 800 W studio strobe (Photon Europe MSN-800) with a white reflective umbrella used as a light diffuser (Photon Europe, 109 cm diameter) mounted onto a 175 cm high light stand, tilted 10° downwards toward the booth. Correct lighting exposure was checked before each session with a digital light meter (Sekonic L-308S). For further details on the photograph acquisition procedure, see Třebický et al. (2018).

Stimuli Processing and Building 360° Head Rotations

Final sets of 36 photographs of full 360° head rotation for each participant were selected and postprocessed in Adobe Lightroom Classic CC (Version 2017, Adobe Systems Inc.). First, we converted photographs into DNG raw files, then we built DNG color calibration profiles and applied them to all photographs. Exposure across all selected photographs was verified in three background areas around the head (above, left, right) and eventual slight differences in exposure were manually adjusted to the same level. Subsequently, the calibrated photographs were exported into lossless 16-bit AdobeRGB TIFF files in real size of

35 × 53 cm and 168 PPI. This allowed us to present images of participants' heads in their real-life size.

Photographs were aligned so that each participant's head was positioned in the center of each frame with eyes on the same horizontal line in all pictures. Final photographs were batch-cropped to 2,095 × 2,305 side ratio to fit head rotations of all participants. All photographs were subsequently converted into sRGB color space and exported as 8-bit JPEG files (2,095 × 2,305 px @ 168 ppi).

We built the 360° head rotations with Sirv (www.sirv.com, Magic Toolbox Limited), an online suite for creating and managing image spins. Photographs of all target participants were uploaded, and individual spins created.

Rating Sessions

Rating sessions took place in a quiet perception lab under standardized conditions. Raters were seated 125 cm from the screen, i.e., at the distance at which the photographs were captured, so as to approximate a social interpersonal distance (Sorokowska et al., 2017) and thereby increase the ecological validity of the rating session.

Ratings were carried out on a 27" Dell U2718Q UltraSharp IPS color calibrated screen (3,840 × 2,160 px, 99% sRGB color space coverage) turned into a vertical position to accommodate the life-sized images used. The data were collected via Qualtrics survey suite (Qualtrics, Provo, UT, United States).

Raters were asked to rate fighting ability ("Jak moc by byl tento muž úspěšný, kdyby se dostal do fyzického souboje?"/"If this man were involved in a physical confrontation, how successful would he be?") of each photograph on a 7-point verbally anchored scale (from "1—velice neúspěšný"/"very unsuccessful," to "7—velice úspěšný"/"very successful"). The 360° rotations spun automatically once and then raters could freely turn the photographs around for further inspection by dragging the mouse left or right before rating them. Photographs were presented in a randomized order and time spent rating was not restricted. Finally, all raters completed a brief questionnaire (regarding their age, height, weight, and self-rated formidability).

Physical Performance and Body Composition Measurements

To determine the physical performance and body parameters of participating athletes, we employed complex measurements relevant to martial arts performance, which included quantifications of their body composition, maximal isometric strength, lung capacity, and AC measurements (Schick et al., 2010; Vidal Andreato et al., 2011; Lenetsky and Harris, 2012; Alm and Yu, 2013; Coufalová et al., 2014; Marinho et al., 2016). **Table 1** provides descriptive statistics. All measurements were performed at the Biomedicine Laboratory of the Faculty of Physical Education and Sport, Charles University (see **Figure 2**).

Body Composition Measurements

To acquire detailed measures of body composition, we performed a bioelectrical impedance analysis, which is based on measuring body's electrical resistance to an imperceptible electric current. Electrical resistance is a function of both body shape and

TABLE 1 | Descriptive statistics of the target sample.

Descriptive statistics				
Characteristic	Mean	SD	Minimum	Maximum
Age (yrs)	26.73	5.91	18	38
Height (cm)	179.82	6.93	165	193.8
Body Weight (kg)	81.47	11.3	60.6	112.4
Body Fat (kg)	7.56	4.53	2.5	21.7
Muscle Mass (kg)	70.27	7.68	55	90
Bone Mass (kg)	3.64	0.37	2.9	4.6
Total Body Water (%)	54.03	6.41	43.2	72.9
Handgrip Strength (right) (kp)	56.27	7.89	39.1	77.4
Handgrip Strength (left) (kp)	54.25	7.43	33.2	77.7
Handgrip Strength (mean) (kp)	55.26	7.33	36.15	77.55
Arm Flexion (right) (kp)	34.5	7.01	22.5	60.9
Arm Flexion (left) (kp)	32.69	7.33	20.5	64.4
Arm Flexion (mean) (kp)	33.59	7.05	22.4	62.65
Arm Extension (right) (kp)	28.41	6.12	17.3	54.3
Arm Extension (left) (kp)	28.69	6.26	10.2	51.3
Arm Extension (mean) (kp)	28.55	6.01	13.75	52.8
Trunk Bend (kp)	94.2	18.36	53.7	166.6
Trunk Forward (kp)	80.46	16.6	49.1	141.8
Neck Forward (kp)	24.03	5.12	14.1	33.7
Neck Bend (kp)	40.26	6.01	31.3	57
Knee Flexion (right) (kp)	31.7	8.27	19.8	57.4
Knee Flexion (left) (kp)	30.17	7.4	19.1	53.4
Knee Flexion (mean) (kp)	30.93	7.7	19.6	55.35
Knee Extension (right) (kp)	69.89	16.91	39.7	116.9
Knee Extension (left) (kp)	64.99	13.26	39.6	95.5
Knee Extension (mean) (kp)	67.44	14.56	39.65	105.35
Forced Vital Capacity (l)	5.28	0.73	4.03	6.84
Forced Expiratory Volume (l)	4.59	0.57	3.13	5.66
Peak Expiratory Flow (l/s)	9.55	1.6	6.66	14.28
Maximum Anaerobic Performance (W)	653.28	130.99	422	966
Minimum Anaerobic Performance (W)	364.76	57.94	197.9	514.1
Average Anaerobic Performance (W)	506.7	87.86	293.1	712
Anaerobic Capacity (kJ)	15.2	2.64	8.8	21.4
Decrease of Performance (W)	290.63	100.83	102.2	521.6
Number of Rotations	46.49	5.77	29	57.1
Actual Fighting Ability (wins/fights ratio)	0.65	0.25	0	1
Mean Fighting Ability Rating	4.05	1.01	2.22	6

the volume of conductive tissues in the body (Goran, 1998). Participating athletes were asked to avoid activities which may bias the measurement, such as consumption of alcoholic beverages, sauna, and demanding physical activities 24 h prior to the test, and eating and drinking for 2 h before the measurement (Brodie et al., 1991; Fogelholm et al., 1993). Body weight,



FIGURE 2 | Physical performance measurements. **Top left**—maximal isometric strength (arm extension dynamometry) measurement, **top right**—lung capacity (spirometry) measurement, **bottom**—anaerobic capacity measurement (Wingate test). Photographs by Jitka Fialová, with informed and written consent of the depicted participant.

body fat mass, muscle mass, and bone mass were measured (Vaara et al., 2012) using Tanita MC-980 bio impedance scale (Athlete setting). Testing was performed in a standing position, with participants both standing on and holding in their hands measuring electrodes with arms freely alongside the body. Participants were tested while wearing only the underwear we provided them with (Pinilla et al., 1992).

Maximal Isometric Strength Measurements

Isometric strength in flexion and extension of arms, legs, trunk, and neck was measured as the peak force produced by maximal voluntary isometric contraction of each muscle group while the athlete was seated on a specifically designed dynamometric station with low profile aluminum load cell (model 1042, measurement error $\pm 0.05\%$) (Coufalová et al., 2014). Using a digital hand-grip dynamometer Takei TKK 5401, we evaluated the isometric strength of hands (Vidal Andreato et al., 2011; Bonitch-Góngora et al., 2013). While performing the hand grip measurements, athletes were instructed to stand straight with their arms alongside their body.

Three attempts were performed for each type of measurement while switching sides between attempts and using the “best test” method, meaning that only the highest performance was recorded and included in subsequent analyses.

Lung Capacity Measurements

To assess the lung capacity, we used spirometry. This physiological test measures how individuals inhale and exhale volumes of air as a function of time while measuring either total volume or flow. Measures of lung capacity were acquired with spirometer MicroLab ML3500 MK8. Three standing forced vital capacity (FVC) maneuvers were performed: we measured the highest volume of FVC, forced expiratory volume in the first second (FEV1), and peak expiratory flow (PEF), while again applying the “best test” method, i.e., recording the highest of three test values. FVC is the maximal volume of air exhaled with maximally forced effort from a maximal inspiration delivered during an expiration made as forcefully and completely as possible (i.e., vital capacity performed with maximally forced expiratory effort). FEV1 is the maximal volume of air exhaled in the first second of forced expiration from a position of full inspiration and PEF indicates the maximum expiratory flow achieved from maximum forced expiration from the point of maximal lung inflation (Miller et al., 2005).

Anaerobic Capacity Measurements

Anaerobic performance we measured using the Wingate test, which consists of 30 s of supramaximal arm-cranking exercise at maximal speed against a frictional resistance determined relative to the subject's body weight (Bar-Or, 1987). Three indices are measured: (1) anaerobic power (AP), which indicates the highest mechanical power elicited during the test, (2) mean power, which shows the average power sustained throughout the 30 s period, (3) AC, which indicates the total work performed during the entire 30 s period, and (4) power decrease (PD), which measures the degree of power drop-off during the test (Collomp et al., 1991). The Wingate test was performed on a Monark arm ergometer (model Rump-Rokos 4.00/C01) with a load of 4 W per kilogram of body weight. Participants were instructed to remain seated and verbally encouraged to perform as quickly as possible right from the start and to maintain maximal turning rates throughout the 30 s period. The test was preceded by a short warm up period, where the participant exercised until achieving 120 bpm heart rate. This was followed by activation of the load (Franchini et al., 2003).

Level of Beardedness

Two authors coded each target's image. To assess the level of facial hair, we employed three beardedness categories defined in earlier research (Dixon et al., 2018). Agreement between both authors was above 95%, i.e., in 42 out of 44 cases; the remaining two cases were discussed and categorized. The procedure resulted in categories: (1) “Shaved” including athletes with no facial hair of any kind ($N = 15$; 34%); (2) “Some beard” including athletes with all kinds of facial hair except shaven and full beards ($N = 20$; 45.5%); (3) “Full beard” including athletes with trimmed and bushy full beards ($N = 9$; 20.5%).

Statistical Analysis

All statistical tests were performed in SPSS 23 (IBM Corp., 2015) and JASP 0.9.0.1 (JASP Team, 2018). McDonald's ω statistics was used to estimate inter-rater agreement. Differences in fighting ability ratings were analyzed by independent samples *t*-test and association between ratings was assessed by bivariate correlations using Pearson's correlation coefficient. Component scores of physical performance measures, which were then used in subsequent analyses, were calculated using principal component analysis (PCA) with no rotations. To assess which factors predict the perceived fighting ability and to estimate their relative contribution, we ran a linear regression analysis where all predictors were entered simultaneously using the enter method. Similarly, we used regression analysis to investigate the relationship between the actual and perceived fighting ability. To test the influence of beardedness on fighting ability ratings, one-way ANOVA was carried out. We entered fighting ability ratings as dependent and level of beardedness as independent variables. The effect size for one-way ANOVA is reported in η_p^2 . A Holm's *post-hoc* test was performed and effect sizes for the comparison are reported in Cohen's *d*.

Component Scores of Physical Performance Measures

In order to reduce the number of variables produced by physical measurements and body composition tests and to obtain robust and representative component scores to apply in subsequent analyses, we used a PCA. We checked the assumptions of this analysis by looking for multicollinearity (>0.9) or singularity ($=0.0$) between variables by a bivariate correlation. For body composition measures, we found a high correlation between body weight, body fat, muscle mass, bone mass, and total body water ($r_s > 0.817$). For later regression analysis, we have therefore decided to keep body weight as the most representative variable that includes all body composition measures. It is a frequently used measure of body size, thus allowing for a comparison with previous research. Analysis of AC data yielded by the Wingate test measurements revealed a high correlation between maximum performance, average performance, AC, and decrease of performance ($r_s > 0.9$). In view of these results, and because we use maximal performance also in other measurements, we decided to use maximum performance as a variable in the PCA. After these initial adjustments, assumptions of the analysis were met.

We subjected maximal isometric strength measurements to the PCA. This produced a single component which we labeled "Isometric strength." Next, we entered spirometry test measurements into the PCA, which resulted in a component we labeled "Lung capacity." Anaerobic capacity measurements also yielded a single component, the "Anaerobic capacity." The resulting components and their loadings are listed in the **Supplementary Materials, Table S1**.

Data Availability

Datasets generated and analyzed during the current study are available in the **Supplementary Material** of this article (Dataset athletes.XLSX, Dataset rating.XLSX).

RESULTS

McDonald's ω scores of ratings by males ($\omega = 0.851$), females ($\omega = 0.738$), and total ($\omega = 0.795$) showed a high inter-rater agreement. In subsequent analyses, we have therefore used mean fighting ability ratings. We have also found a high correlation between fighting ability ratings assigned by men and women ($r = 0.972$, 95% CI [0.95, 0.985], $p < 0.001$), which is why we decided to analyze the ratings of both sexes together. Independent samples *t*-test also showed no significant sex differences in ratings [$t_{(86)} = 0.041$, $p = 0.968$, $d = 0.009$], which further supported our decision to analyze the ratings of both sexes together.

Predictors of Perceived Fighting Ability

A multiple linear regression analysis was run to predict perceived fighting ability whereby age, weight, Isometric strength, Lung capacity, and AC components were all treated as independent predictors. The overall model was significant [$F_{(5, 38)} = 2.79$, $p = 0.031$, $R^2 = 0.269$], but none of the individual predictors statistically significantly predicted the perception of fighting ability: all $p_s > 0.05$ (see **Table 2**).

Actual Fighting Ability as a Predictor of Perceived Fighting Ability

Exploratory correlation analysis showed a positive correlation between fighter's age ($r = 0.35$, $p = 0.018$), weight ($r = 0.341$, $p = 0.022$), and perceived fighting ability, which is why we added these measures into the linear regression model. The overall model significantly predicted perceived fighting ability [$F_{(3, 40)} = 3.579$, $p = 0.022$, $R^2 = 0.212$]. Among the predictors, body weight significantly predicted perceived fighting ability ($\beta = 0.31$, $t = 2.033$, $p = 0.049$), but actual fighting ability ($\beta = -0.175$, $t = -1.205$, $p = 0.235$) nor age ($\beta = 0.247$, $t = 1.669$, $p = 0.103$) were statistically significantly related to perceived fighting ability (see **Table 3**).

The Effect of Beardedness on Perceived Fighting Ability

We found a moderate effect bordering on a formal level of significance of beardedness (Shaved: $M = 3.55$, $SD = 1.09$; Some beard: $M = 4.3$, $SD = 0.8$; Full beard: $M = 4.34$, $SD = 1.07$) on fighting ability rating [$F_{(2, 41)} = 3.099$, $p = 0.056$, $\eta_p^2 = 0.131$]. For exploratory purposes, we ran Holm's *post-hoc* comparison. Although not significantly, the Shaved category received the lowest rating, while Some beard ($t = 2.279$, $p_{Holm} = 0.084$, Cohen's $d = 0.801$) and Full beard ($t = 1.943$, $p_{Holm} = 0.118$, Cohen's $d = 0.729$) categories received higher ratings. Some beard and Full beard categories did not differ ($t = 0.102$, $p_{Holm} = 0.919$, Cohen's $d = 0.044$) (**Figure 3**).

DISCUSSION

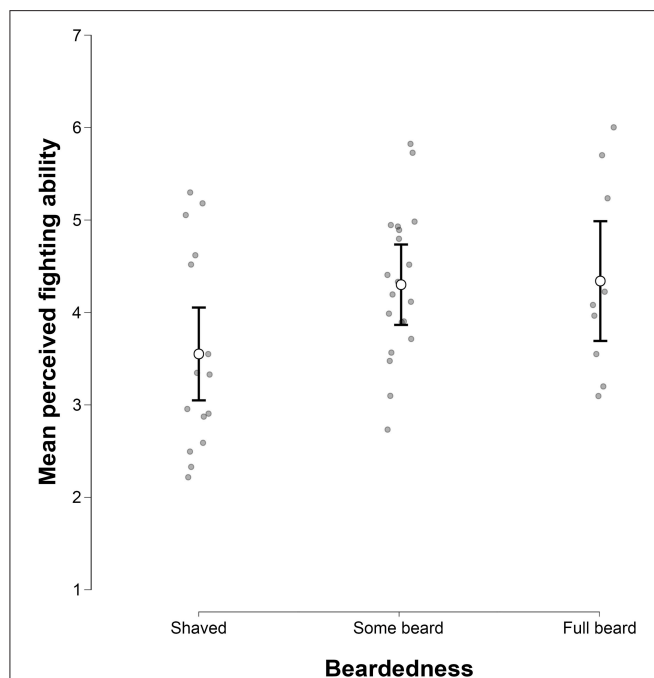
In the present study, we tested which aspects of physical performance affect the perception of fighting ability. To that purpose, we used 360° head rotation photographs of male MMA athletes. We gathered detailed physical measures relevant to physical confrontations, and although it turned out that

TABLE 2 | A summary of regression analysis for variables predicting the perceived fighting ability (fighters' age, body weight, Isometric strength, Lung capacity, and Anaerobic capacity component).

Variable	Unstandardized coefficients		Standardized coefficients	<i>t</i>	<i>p</i>	95 % CI for B		Correlations		
	B	SE				Lower	Upper	Zero-order	Partial	Part
(Constant)	0.931	1.542		0.604	0.549	−2.190	4.052			
Age	0.032	0.029	0.186	1.097	0.279	−0.027	0.090	0.347	0.175	0.152
Body weight	0.028	0.018	0.313	1.572	0.124	−0.008	0.064	0.345	0.247	0.218
Isometric strength component	−0.284	0.181	−0.282	−1.569	0.125	−0.651	0.083	0.095	−0.247	−0.218
Lung capacity component	0.003	0.162	0.003	0.018	0.985	−0.324	0.330	0.098	0.003	0.003
Anaerobic capacity component	0.296	0.182	0.294	1.631	0.111	−0.071	0.664	0.418	0.256	0.226

TABLE 3 | Summary of regression analysis for the relationship between perceived and actual fighting ability, fighters' age, and body weight.

Variable	Unstandardized coefficients		Standardized coefficients	<i>t</i>	<i>p</i>	95% CI for B		Correlations		
	B	SE				Lower	Upper	Zero-order	Partial	Part
(Constant)	1.141	1.099		1.038	0.306	−1.081	3.363			
Actual fighting ability (wins/fights ratio)	−0.715	0.594	−0.175	−1.205	0.235	−1.915	0.485	−0.109	−0.187	−0.169
Age	0.042	0.025	0.247	1.669	0.103	−0.009	0.093	0.347	0.255	0.234
Body weight	0.028	0.014	0.310	2.033	0.049	0.000	0.055	0.345	0.306	0.285

**FIGURE 3** | Differences in mean ratings of fighting ability between levels of beardedness (Shaved, Some beard, Full beard). The graph represents the means, their 95% CIs, and data distribution for the three beardedness levels. Mean perceived fighting ability did not differ significantly between beardedness levels.

overall physical performance predicts fighting ability rating, statistical analysis did not show that any particular predictor contributes to the perception of fighting ability significantly.

Body weight and AC (AC component) did, however, explained most of the variability, while isometric strength (Isometric strength component) was related negatively. We did not find any significant association between the perceived fighting ability and actual fighting ability in physical confrontations, and perceived fighting ability was predicted solely by athletes' body weight. Further, we explored a possible effect of beardedness on perceived fighting ability and found a moderate-sized but non-significant effect, whereby the shaved targets received the lowest rating.

Compared to previous investigations which used either just one or a limited number of threat potential measures, we collected detailed data about various aspects of physical performance. Although other studies (e.g., Sell et al., 2009; Han et al., 2017) have reported that handgrip strength, height, and weight affect the perception of overall fighting ability, our aim was to investigate other relevant factors which could potentially contribute to perceptual inferences, such as overall isometric strength, anaerobic, or lung capacity. Our regression model significantly predicted perceived fighting ability but none of the individual predictors contributed to the perceived fighting ability significantly. We identified body weight and AC as variables that have the greatest impact on perceived fighting ability. The general probability of being perceived as a winner seems to be related to body size and weight (Deaner et al., 2012), whereby heavier athletes are seen as better fighters than the lighter ones.

It has been suggested in earlier studies that body size (here assessed as body weight) plays a key role during the initial phase of formidability assessments (Třebický et al., 2015). We could thus speculate that our findings are compatible with a model according to which assessment of a potential opponent takes place on multiple levels (Třebický and Havlíček, 2017). The first step, the "fight or flight" decision, seems to depend

mainly on opponent's overall size. If, however, the rivals are of a comparable size, another level of assessment may be deployed that could be linked to the perception of other potentially significant characteristics.

Another predictor of perceived fighting ability was AC as measured by the Wingate test. Anaerobic capability has been reported as a key characteristic of successful martial arts athletes (James et al., 2016). MMA is physiologically complex and during contests, fighters deploy a wide range of mechanical and metabolic qualities. Intense striking exchanges are common, but twice as many fights end during highly physically demanding ground fight sequences, that is, when fighters use their wrestling and grappling techniques (Del Vecchio et al., 2011). High-intensity and relatively long engagements are therefore a significant part of the overall performance, and they can be approximated by the Wingate test. Moreover, this measure seems to be related to general physical fitness and performance, which is apparent from its correlation with body composition, isometric strength and lung capacity (see exploratory correlations **Table S2** in Supplementary Materials). Earlier research has revealed that cues to strength are present in human faces (Hugill et al., 2009) and individuals' strength is connected to masculinity and dominance ratings (Fink et al., 2007; Windhager et al., 2011), but we found no evidence of a similar relationship for assessments of fighting ability. In fact, our data suggest a rather surprising opposite pattern. Although physical performance is undoubtedly the cornerstone of successful performance, psychological characteristics such as personality, ability to cope with stress, or successful execution of techniques and other skills may also significantly affect fighting ability and success (Gould et al., 1981; Filaire et al., 2001; Radochonski et al., 2011; Ruiz and Hanin, 2011; Chen and Cheesman, 2013; Bernacka et al., 2016). These factors, however, exceed the scope of the current study and should be investigated in future.

Our results are also in contrast with former studies which showed that people can accurately assess actual fighting ability from facial photographs of MMA fighters when asked to rate their aggressiveness, fighting ability, or likelihood of winning (Třebický et al., 2013; Little et al., 2015). Třebický et al. (2013) found a link between the perceived and actual fighting ability only in heavyweight fighters. Limited sample size and uneven distribution within weight categories did not allow us to test the effect of weight category directly, but heavier athletes in our sample were perceived as more formidable competitors, which suggests a similar pattern (see **Table S2**).

Numerous studies have shown that beards enhance ratings of traits related to intrasexual competition, such as men's perceived age, masculinity, social dominance, or aggressiveness. In short, bearded men tend to score higher in these measures than clean-shaved individuals (Neave and Shields, 2008; Dixon and Vasey, 2012; Geniole and McCormick, 2015; Saxton et al., 2016). Our study provides additional support to these findings because athletes with facial hair were rated higher on the perceived fighting ability scale.

The main goal of present study was to investigate the visual perception of threat potential. We have therefore employed a more holistic concept of fighting ability and did not focus

solely on the perception of particular characteristics which may contribute to fighting success. In other studies, participants were asked to assess strength, dominance, masculinity, and aggressiveness, which are all relatively simple characteristics. In the current study, participants rated fighting ability, an arguably more abstract or comprehensive quality, which made the assessments more difficult to process. One could speculate that the use of a different rating scale, e.g., one focused on aggressiveness, could yield significant findings, because earlier studies (e.g., Třebický et al., 2013) have reported a close association between this characteristic and perceived fighting ability.

Future studies should address also non-European populations of raters (Třebický et al., 2018). It is possible that other aspects of the male physique may be associated with perceived fighting ability for instance in Asian cultures, where agility, flexibility, and movement complexity may play a more important role. It is also possible that originally African fighting styles were more dynamic than the rather static and force-oriented European styles.

Our sample substantially differs from previous studies in several aspects. Athletes in the present study varied in performance levels, ranging all the way from beginner amateurs to seasoned professionals, while earlier studies used as stimuli photographs of high-profile professional fighters (UFC). Professional fighters have a considerably greater fighting record, which translates into more accurate estimation of their actual fighting ability. In our study, we included fighters who took part in at least two fights, but this low number of fights may result in an inaccurate picture of athlete's true fighting potential, especially in case of fighters who are just starting their careers. Moreover, reliability of the score could be affected by the way in which fighters are paired for matches, which is not a random process. Fighters are paired by organizing committees who take into account their previous experience and fighting record. This may potentially limit the use of the success score as a measure of actual fighting ability. One could think of a more complex measure of fighting ability which would take into consideration the formidability of the opponent (e.g., winning a fight against an experienced fighter would result in a bonus score, i.e., a higher score than winning a fight with a beginner). Data on formidability/experience of the individual fighters were not, however, obtained in our sample and the wins-to-all-fights ratio remains the most objective measure of the fighting ability available at the moment.

The present study is also based on a relatively small sample and one could therefore argue that it had a rather limited chance of detecting expected associations. Nevertheless, our sample size is comparable to earlier studies such as Han et al. (2017) ($N = 44$); Sell et al. (2009) ($N = 59$) or Windhager et al. (2011) ($N = 26$). Despite our best efforts, we found no more volunteers among fighters who would meet our criteria (age 18–40, at least two MMA fights) and were willing to participate, because the popularity of MMA in the Czech Republic is increasing rather slowly.

To limit a potential systematic bias in photographs and to give our raters maximum visual information, we took highly standardized 360° rotating photographs. We have also

asked participants to adopt a neutral expression and straight position during the whole photo session, which could help eliminate possible cues to fighting ability inferences based on slight and unintentional facial expressions. This is an important point because several earlier studies used downloaded photographs of professional fighters, which varied in lighting conditions, head tilts etc. Variability in stimuli quality in previous research could be viewed as “noise” which decreased the likelihood of finding a systematic effect. Nonetheless, can we really take it for granted that this assumption is correct? Could it be that a more formidable depiction, e.g., one with a head tilt and a frown, is a cue to better fighters? The effect of image standardization vs. self-expression on accuracy of inferences should be addressed in future studies.

One may also argue that the generalization potential of results based solely on MMA fighter population may be limited due to its specific characteristics, such as overall high level of formidability and rather specific appearance (broken noses, facial scars, etc.). We tried to decrease potential bias by informing the raters about the target selection criteria only upon completion of the study. Interestingly, the mean formidability rating on a 7-point scale was ≈ 4 (ranging from 2.22 to 6) and the data followed a normal distribution. Although the physical performance of the MMA fighters may be considerably higher than that found in the general population in industrialized countries, it may be less impressive when compared to aged-matched individuals from small-scale societies. It is thus possible that a high level of athletic performance mirrors ancestral human conditions better than the commonly used student samples.

In conclusion, we found no significant connections between the measured predictors of physical performance and the perception of fighting ability from facial photographs. Based on observed effect sizes, we can tentatively conclude that inferences of fighting ability are mostly linked to body size (especially weight) and AC, which are both qualities which affect the outcome of physical confrontations. Our results therefore indicate that the perception of fighting ability may be more complex than previously thought.

REFERENCES

- Alm, P., and Yu, J. (2013). Physiological characters in mixed martial arts. *Am. J. Sports Sci.* 1, 12–17. doi: 10.11648/j.jss.20130102.11
- Baldassare, M., and Feller, S. (1975). Cultural variations in personal space: theory, methods, and evidence. *Ethos* 3, 481–503. doi: 10.1525/eth.1975.3.4.02a00020
- Bar-Or, O. (1987). The Wingate anaerobic test an update on methodology, reliability and validity. *Sports Med.* 4, 381–394. doi: 10.2165/00007256-198704060-00001
- Bernacka, R. E., Sawicki, B., Mazurek-Kusiak, A., and Hawlena, J. (2016). Conforming and nonconforming personality and stress coping styles in combat athletes. *J. Hum. Kinet.* 51, 225–233. doi: 10.1515/hukin-2015-0186
- Boileau, R. A., and Lohman, T. G. (1977). The measurement of human physique and its effect on physical performance. *Orthop. Clin. North Am.* 8, 563–581.
- Bonitch-Góngora, J., Almeida, F., Padial, P., Bonitch-Domínguez, J., and Feriche, B. (2013). Maximal isometric handgrip strength and endurance differences between elite non-elite young judo athletes. *Arch. Budo* 9, 239–248.

AUTHOR CONTRIBUTIONS

VT, JF, and JH developed the study concept. KK and DS contributed to the study design. Data collection was performed by VT, JF, DS, KC, RK, ZŠ, and RP. VT and JF performed the data analysis and interpretation. VT, JF, and JH drafted the manuscript. DS, KC, KK, RK, ZŠ, and RP provided critical revisions. All authors approved the final version of the manuscript for submission.

FUNDING

This research was supported by Czech Science Foundation GACR P407/16/03899S, by the Charles University Research Centres UNCE/HUM/032, UNCE 204056, and by the Ministry of Education, Youth, and Sports NPU I program (No. LO1611).

ACKNOWLEDGMENTS

We thank Tereza Nevolová, Žaneta Slámová, Dagmar Schwambergová, Pavel Šebesta, and other members of Human Ethology group (www.etologiecloveka.cz) for their help with data collection and ratings, Zsófia Csajbók for valuable advices on statistical analysis, Mixed Martial Arts Association Czech Republic (MMAA), Anna Pilátová, Ph.D. for proofreading, and all volunteers for their participation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02740/full#supplementary-material>

Video 1 | Sample of 360° head rotation.

Table S1 | Factor analysis; Exploratory correlation matrix.

Table S2 | Dataset for formidability rating.

Table S3 | Dataset for Athletes measurements.

- Braswell, M. T., Szymanski, D. J., Szymanski, J. M., Dixon, E. E., Gilliam, S. T., Wood, R. J., et al. (2010). Physiological differences in mixed martial artist and traditional martial artists: a pilot study. *J. Strength Cond. Res.* 24, 1. doi: 10.1097/01.JSC.0000367074.45565.85
- Brodie, D. A., Eston, R. G., Coxon, A. Y., Kreitzman, S. N., Stockdale, H. R., and Howard, A. N. (1991). Effect of changes of water and electrolytes on the validity of conventional methods of measuring fat-free mass. *Ann. Nutri. Metabol.* 35, 89–97. doi: 10.1159/000177631
- Bryan, R., Perona, P., and Adolphs, R. (2012). Perspective distortion from interpersonal distance is an implicit visual cue for social judgments of faces. *PLoS ONE* 7:e45301. doi: 10.1371/journal.pone.0045301
- Chen, M. A., and Cheesman, D. J. (2013). Mental toughness of mixed martial arts athletes at different levels of competition. *Percept. Mot. Skills* 116, 905–917. doi: 10.2466/29.30.PMS.116.3.905-917
- Collomp, K., Ahmaidi, S., Audran, M., Chanal, J.-L., and Préfaut, C. (1991). Effects of caffeine ingestion on performance and anaerobic metabolism during the Wingate test. *Int. J. Sports Med.* 12, 439–443. doi: 10.1055/s-2007-1024710

- Coufalová, K., Cochrane, D. J., Malý, T., and Heller, J. (2014). Changes in body composition, anthropometric indicators and maximal strength due to weight reduction in judo. *Arch. Budo* 10, 161–168.
- Danel, D. P., Valentova, J. V., Sanchez, O. R., Leongomez, J. D., Varella, M. A. C., and Kleisner, K. (2018). A cross-cultural study of sex-typicality and averageness: correlation between frontal and lateral measures of human faces. *Am. J. Hum. Biol.* 30:e23147. doi: 10.1002/ajhb.23147
- Deaner, R. O., Goetz, S. M. M., Shattuck, K., and Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *J. Res. Per.* 46, 235–238. doi: 10.1016/j.jrp.2012.01.005
- Del Vecchio, F. B., Hirata, S. M., and Franchini, E. (2011). A review of time-motion analysis and combat development in Mixed Martial Arts matches at regional level tournaments. *Percept. Motor Skills* 112, 639–648. doi: 10.2466/05.25.PMS.112.2.639-648
- Dixon, B. J., and Vasey, P. L. (2012). Beards augment perceptions of men's age, social status, and aggressiveness, but not attractiveness. *Behav. Ecol.* 23, 481–490. doi: 10.1093/beheco/arr214
- Dixon, B. J. W., Lee, A. J., Sherlock, J. M., and Talamas, S. N. (2017). Beneath the beard: do facial morphometrics influence the strength of judgments of men's beardedness? *Evol. Hum. Behav.* 38, 164–174. doi: 10.1016/j.evolhumbehav.2016.08.004
- Dixon, B. J. W., Sherlock, J. M., Cornwell, W. K., and Kasumovic, M. M. (2018). Contest competition and men's facial hair: beards may not provide advantages in combat. *Evol. Hum. Behav.* 39, 147–153. doi: 10.1016/j.evolhumbehav.2017.11.004
- Durmic, T., Lazovic Popovic, B., Zlatkovic Svenda, M., Djelic, M., Zugic, V., Gavrilovic, T., et al. (2017). The training type influence on male elite athletes' ventilatory function. *BMJ Open Sport Exerc. Med.* 3:e000240. doi: 10.1136/bmjsem-2017-000240
- Ellis, L. (1995). Dominance and reproductive success among nonhuman animals: a cross-species comparison. *Ethol. Sociobiol.* 16, 257–333. doi: 10.1016/0162-3095(95)00050-U
- Erkelens, C. J. (2018). Multiple pictures of a perspective scene reveal the principles of picture perception. *Vision* 2, 1–17. doi: 10.3390/vision2030026
- Filaire, E., Maso, F., Sagnol, M., Ferrand, C., and Lac, G. (2001). Anxiety, hormonal responses, and coping during a judo competition. *Aggress. Behav.* 27, 55–63. doi: 10.1002/1098-2337(20010101/31)27:1<55::AID-AB5>3.0.CO;2-H
- Fink, B., Neave, N., and Seydel, H. (2007). Male facial appearance signals physical strength to women. *Am. J. Hum. Biol.* 19, 82–87. doi: 10.1002/ajhb.20583
- Fogelholm, M., Sievänen, H., Kukkonen-Harjula, K., Oja, P., and Vuori, I. (1993). "Effects of meal and its electrolytes on bioelectrical impedance," in *Human Body Composition*, eds K. Ellis and J. Eastman (Boston, MA: Springer), 331–332. doi: 10.1007/978-1-4899-1268-8_75
- Franchini, E., Yuri Takito, M., Yuzo Nakamura, F., Ayumi Matsushigue, K., and Peduti Dal'Molin Kiss, M. A. (2003). Effects of recovery type after a judo combat on blood lactate and performance in an intermittent anaerobic tests. *J. Sports Med. Phys. Fitness* 43, 424–431. doi: 10.1007/S00421-009-1134-2
- Gallup, A. C., and Fink, B. (2018). Handgrip strength as a darwinian fitness indicator in men. *Front. Psychol.* 9:439. doi: 10.3389/fpsyg.2018.00439
- Geniole, S. N., and McCormick, C. M. (2015). Facing our ancestors: judgements of aggression are consistent and related to the facial width-to-height ratio in men irrespective of beards. *Evol. Hum. Behav.* 36, 279–285. doi: 10.1016/j.evolhumbehav.2014.12.005
- Goran, M. I. (1998). Measurement issues related to studies of childhood obesity: assessment of body composition, body fat distribution, physical activity, and food intake. *Am. Acad. Pediatr.* 101, 505–518.
- Gould, D., Weiss, M., and Weinberg, R. (1981). Psychological characteristics of successful and unsuccessful big ten wrestlers. *J. Sport Psychol.* 3, 69–81. doi: 10.1123/jsp.3.1.69
- Hall, E. T. (1966). *The Hidden Dimension*. New York, NY: Doubleday & Co.
- Han, C., Kandrik, M., Hahn, A. C., Fisher, C. I., Feinberg, D. R., Holzeitner, I. J., et al. (2017). Interrelationships among men's threat potential, facial dominance, and vocal dominance. *Evol. Psychol.* 15:474704917697332. doi: 10.1177/1474704917697332
- Hill, A. K., Hunt, J., Welling, L. L. M., Cárdenas, R. A., Rotella, M. A., Wheatley, J. R., et al. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evol. Hum. Behav.* 34, 334–341. doi: 10.1016/j.evolhumbehav.2013.05.004
- Holzeitner, I. J., and Perrett, D. I. (2016). Perception of strength from 3D faces is linked to facial cues of physique. *Evol. Hum. Behav.* 37, 217–229. doi: 10.1016/j.evolhumbehav.2015.11.004
- Hugill, N., Fink, B., Neave, N., and Seydel, H. (2009). Men's physical strength is associated with women's perceptions of their dancing ability. *Personal. Individ. Differ.* 47, 527–530. doi: 10.1016/j.paid.2009.04.009
- IBM Corp. (2015). *IBM SPSS Statistics for Windows, Version 23.0*. Armonk, NY: IBM Corp.
- James, L. P., Haff, G. G., Kelly, V. G., and Beckman, E. M. (2016). Towards a determination of the physiological characteristics distinguishing successful mixed martial arts athletes: a systematic review of combat sport literature. *Sports Med.* 46, 1525–1551. doi: 10.1007/s.40279-016-0493-1
- JASP Team (2018). *JASP (Version 0.9.0.1)*. Available online at: <https://jasp-stats.org/>
- La Bountry, P., Campbell, B. I., Galvan, E., Cooke, M., and Antonio, J. (2011). Strength and conditioning considerations for mixed martial arts. *Strength Cond. J.* 33, 56–67. doi: 10.1519/SSC.0b013e3182044304
- Lenetsky, S., and Harris, N. (2012). The mixed martial arts athlete: a physiological profile. *Strength Cond. J.* 34, 32–47. doi: 10.1519/SSC.0b013e3182389f00
- Little, A. C., Třebický, V., Havlíček, J., Roberts, S. C., and Kleisner, K. (2015). Human perception of fighting ability: facial cues predict winners and losers in mixed martial arts fights. *Behav. Ecol.* 26, 1470–1475. doi: 10.1093/beheco/arv089
- Marinho, B. F., Follmer, B., Del Conti Esteves, J. V., and Andreato, L. V. (2016). Body composition, somatotype, and physical fitness of mixed martial arts athletes. *Sport Sci. Health* 12, 157–165. doi: 10.1007/s11332-016-0270-4
- Miller, M. R., Hankinson, J., Brusasco, V., Burgos, F., Casaburi, R., Coates, A., et al. (2005). Standardisation of spirometry. *Eur. Resp. J.* 26, 319–338. doi: 10.1183/09031936.05.00034805
- Muscarella, F., and Cunningham, M. R. (1996). The evolutionary significance and social perception of male pattern baldness and facial hair. *Ethol. Sociobiol.* 17, 99–117. doi: 10.1016/0162-3095(95)00130-1
- Neave, N., and Shields, K. (2008). The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces. *Pers. Individ. Dif.* 45, 373–377. doi: 10.1016/j.paid.2008.05.007
- Pinilla, J. C., Webster, B., Baetz, M., Reeder, B., Hattori, S., and Liu, L. (1992). Effect of body positions and splints in bioelectrical impedance analysis. *J. Parent. Enteral. Nutri.* 16, 408–412. doi: 10.1177/0148607192016005408
- Putts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Putts, D. A., Apicella, C. L., and Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proc. R. Soc. B Biol. Sci.* 279, 601–609. doi: 10.1098/rspb.2011.0829
- Radachowski, M., Cynarski, W., Perenc, L., and Siorek-Maślanka, L. (2011). Competitive anxiety and coping strategies in young martial arts and track and field athletes. *J. Hum. Kinet.* 27, 180–189. doi: 10.2478/v10078-011-0014-0
- Radovanovic, D., Bratic, M., Nurkic, M., and Stankovic, N. (2011). Recovery of dynamic lung function in elite judoists after short-term high intensity exercise. *Arch. Budo* 7, 21–26.
- Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J., and Reby, D. (2018). Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* 4, 273–280. doi: 10.1016/j.isci.2018.05.002
- Rowland, H. M., and Burriss, R. P. (2017). Human colour in mate choice and competition. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160350. doi: 10.1098/rstb.2016.0350
- Ruiz, M. C., and Hanin, Y. L. (2011). Perceived impact of anger on performance of skilled karate athletes. *Psychol. Sport Exerc.* 12, 242–249. doi: 10.1016/j.psychsport.2011.01.005
- Saxton, T. K., Mackey, L. L., McCarty, K., and Neave, N. (2016). A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. *Behav. Ecol.* 27, 512–519. doi: 10.1093/beheco/arv178
- Schick, M. G., Brown, L. E., Coburn, J. W., Beam, W. C., Schick, E. E., and Dabbs, N. C. (2010). Physiological profile of mixed martial artists. *Med. Sportiva* 14, 182–187. doi: 10.2478/v10036-010-0029-y
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. B Biol. Sci.* 277, 3509–3518. doi: 10.1098/rspb.2010.0769

- Sell, A., Lukaszewski, A. W., and Townsley, M. (2017). Cues of upper body strength account for most of the variance in men's bodily attractiveness. *Proc. R. Soc. B Biol. Sci.* 284:20171819. doi: 10.1098/rspb.2017.1819
- Sell, A. N. (2016). "Male adaptations to assess fighting ability," in *Encyclopedia of Evolutionary Psychological Science*, eds T. Shackelford and V. Weekes-Shackelford (Cham: Springer). doi: 10.1007/978-3-319-16999-6_937-1
- Sell, A. N., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., and Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proc. R. Soc. B Biol. Sci.* 276, 575–584. doi: 10.1098/rspb.2008.1177
- Sell, A. N., Hone, L. S. E., and Pound, N. (2012). The importance of physical strength to human males. *Hum. Nat.* 23, 30–44. doi: 10.1007/s12110-012-9131-2
- Sherlock, J. M., Tegg, B., Sulikowski, D., and Dixon, B. J. W. (2017). Facial masculinity and beardedness determine men's explicit, but not their implicit, responses to male dominance. *Adapt. Hum. Behav. Physiol.* 3, 14–29. doi: 10.1007/s40750-016-0047-7
- Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., et al. (2017). Preferred interpersonal distances: a global comparison. *J. Cross-Cult. Psychol.* 48, 577–592. doi: 10.1177/0022022117698039
- Thorstenon, C. A. (2018). The social psychophysics of human face color: review and recommendations. *Soc. Cogn.* 36, 247–273. doi: 10.1521/soco.2018.36.2.247
- Třebický, V., Fialová, J., Kleisner, K., and Havlíček, J. (2016). Focal length affects depicted shape and perception of facial images. *PLoS ONE* 11:e0149313. doi: 10.1371/journal.pone.0149313
- Třebický, V., Fialová, J., Kleisner, K., Roberts, S. C., Little, A. C., and Havlíček, J. (2015). Further evidence for links between facial width-to-height ratio and fighting success: commentary on Zilioli et al. (2014). *Aggres. Behav.* 41, 331–334. doi: 10.1002/ab.21559
- Třebický, V., Fialová, J., Stella, D., Štěrbová, Z., Kleisner, K., and Havlíček, J. (2018). 360 degrees of facial perception: congruence in perception of frontal portrait, profile, and rotation photographs. *Front. Psychol.* 9:2405. doi: 10.3389/fpsyg.2018.02405
- Třebický, V., and Havlíček, J. (2017). "Signals of body size," in *Encyclopedia of Evolutionary Psychological Science*, eds T. K. Shackelford and V. A. Weekes-Shackelford (Cham: Springer). doi: 10.1007/978-3-319-16999-6_2739-1
- Třebický, V., Havlíček, J., Roberts, S. C., Little, A. C., and Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. *Psychol. Sci.* 24, 1664–1672. doi: 10.1177/0956797613477117
- Třebický, V., Kleisner, K., and Havlíček, J. (2012). Evolutionary concepts of human physical attractiveness: the case of male physique. *Anthropologie* 50, 33–45. doi: 10.1146/annurev.anthro.33.070203.143733
- Třebický, V., Saribay, S. A., Kleisner, K., Akoko, R. M., Kočnar, T., Valentova, J. V., et al. (2018). Cross-cultural evidence for apparent racial outgroup advantage: congruence between perceived facial aggressiveness and fighting success. *Sci. Rep.* 8:9767. doi: 10.1038/s41598-018-27751-0
- Vaara, J. P., Kyröläinen, H., Niemi, J., Ohrankämmen, O., Häkkinen, A., Kocay, S., et al. (2012). Associations of maximal strength and muscular endurance test scores with cardiorespiratory fitness and body composition. *J. Strength Cond. Res.* 26, 2078–2086. doi: 10.1519/JSC.0b013e31823b06ff
- Vidal Andreato, L., Franzói de Moraes, S. M., Lopes de Moraes Gomes, T., Del Conti Esteves, J. V., Vidal Andreato, T., and Franchini, E. (2011). Estimated aerobic power, muscular strength and flexibility in elite Brazilian Jiu-Jitsu athletes. *Sci. Sports* 26, 329–337. doi: 10.1016/j.scispo.2010.12.015
- von Rueden, C., Gurven, M., and Kaplan, H. (2008). The multiple dimensions of male social status in an Amazonian society. *Evol. Hum. Behav.* 29, 402–415. doi: 10.1016/j.evolhumbehav.2008.05.001
- Walker, P. L. (2001). A bioarchaeological perspective on the history of violence. *Ann. Rev. Anthropol.* 30, 573–596. doi: 10.1146/annurev.anthro.30.1.573
- Wind, A. E., Takken, T., Helder, P. J. M., and Engelbert, R. H. H. (2010). Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *Eur. J. Pediatr.* 169, 281–287. doi: 10.1007/s00431-009-1010-4
- Windhager, S., Schaefer, K., and Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *Am. J. Hum. Biol.* 23, 805–814. doi: 10.1002/ajhb.21219
- Yoon, J. (2002). Physiological profiles of elite senior wrestlers. *Sport. Med.* 32, 225–233. doi: 10.2165/00007256-200232040-00002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Třebický, Fialová, Stella, Coufalová, Pavelka, Kleisner, Kuba, Štěrbová and Havlíček. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Toward a New Approach to Cross-Cultural Distinctiveness and Typicality of Human Faces: The Cross-Group Typicality/Distinctiveness Metric

Karel Kleisner^{1*}, Šimon Pokorný¹ and S. Adil Saribay²

¹ Department of Philosophy and History of Science, Charles University, Prague, Czechia, ² Department of Psychology, Boğaziçi University, Istanbul, Turkey

OPEN ACCESS

Edited by:

Ian Stephen,
Macquarie University, Australia

Reviewed by:

Al-Suan Lee,
Universiti Tunku Abdul Rahman,
Malaysia

Justin Kyle Mogilski,
University of South Carolina
Salkehatchie, United States

Anthony James Lee,
University of Glasgow,
United Kingdom

Iris J. Holzleitner,
University of Glasgow,
United Kingdom

*Correspondence:

Karel Kleisner
karel.kleisner@natur.cuni.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 11 September 2018

Accepted: 14 January 2019

Published: 31 January 2019

Citation:

Kleisner K, Pokorný Š and
Saribay SA (2019) Toward a New
Approach to Cross-Cultural
Distinctiveness and Typicality
of Human Faces: The Cross-Group
Typicality/Distinctiveness Metric.
Front. Psychol. 10:124.
doi: 10.3389/fpsyg.2019.00124

In the present research, we took advantage of geometric morphometrics to propose a data-driven method for estimating the individual degree of facial typicality/distinctiveness for cross-cultural (and other cross-group) comparisons. Looking like a stranger in one's home culture may be somewhat stressful. The same facial appearance, however, might become advantageous within an outgroup population. To address this fit between facial appearance and cultural setting, we propose a simple measure of distinctiveness/typicality based on position of an individual along the axis connecting the facial averages of two populations under comparison. The more distant a face is from its ingroup population mean toward the outgroup mean the more distinct it is (vis-à-vis the ingroup) and the more it resembles the outgroup standards. We compared this new measure with an alternative measure based on distance from outgroup mean. The new measure showed stronger association with rated facial distinctiveness than distance from outgroup mean. Subsequently, we manipulated facial stimuli to reflect different levels of ingroup-outgroup distinctiveness and tested them in one of the target cultures. Perceivers were able to successfully distinguish outgroup from ingroup faces in a two-alternative forced-choice task. There was also some evidence that this task was harder when the two faces were closer along the axis connecting the facial averages from the two cultures. Future directions and potential applications of our proposed approach are discussed.

Keywords: typicality, distinctiveness, geometric morphometrics, cross-culture, face space, morphology

INTRODUCTION

Travelers to a foreign country are sometimes mistaken to be local. One of the authors has discovered on his many trips to Turkey as a Czech that he can easily mislead people to believe that he is from the Black Sea region of Turkey. Another of the current authors, during his sabbatical in Prague as a Turk, has frequently found himself being spoken to in Czech by locals who did not realize he is a foreigner. It is possible that such experiences are partly the result of some level of resemblance of our respective faces to the typical outgroup face.

Typicality and the concept of type has traditionally played an important role within all life sciences including psychology, comparative anatomy, and morphology (Galton, 1879, 1883; Russell, 1916; Goethe, 1999; Kleisner, 2007). Typical object, or abstraction of a typical object, results from comparison of many particular occurrences of things. Such typical objects are usually considered as a reference against which all other things in the environment are evaluated. The things perceived as the most distant from a type are realized as distinct and less familiar.

Despite the extensive interest in facial distinctiveness, measuring distinctiveness is somewhat complicated because numerous facial aspects and their interrelatedness determine whether a face is perceived as distinctive (Wickham et al., 2000). In his influential work, Valentine (1991) defined facial distinctiveness as a function of Euclidean distance from populational mean face. In Valentine's Face Space, the understanding of typical and distinctive faces is not separated and both are covered by the same multidimensional framework. In such multidimensional similarity space, faces are represented as single points in high-dimensional similarity space defined by visual properties (or facial measurements); faces are normally distributed and there is a higher density of faces closer to origin (mean face); typical faces are closer to the origin (mean face) than atypical faces; typical faces are around the mean while atypical faces are on the periphery (Valentine, 1991; Valentine et al., 2016). Valentine's model is, however, limited to intra-population comparisons since outgroup faces will necessarily form a cluster far away from ingroup mean.

Previous research produced evidence that typicality does affect social perception of faces, focusing mostly on the relationship between face typicality and attractiveness (see, e.g., Langlois and Roggman, 1990; Perrett et al., 1994; Rhodes and Tremewan, 1996; Fink and Penton-Voak, 2002; Rhodes, 2006; DeBruine et al., 2007; Said and Todorov, 2011; Danel et al., 2012; Trujillo et al., 2014). Recently, Sofer et al. (2015) demonstrated that face typicality plays an important role in trustworthiness judgments showing that perceived trustworthiness, but not attractiveness, changes along the cline of facial typicality (Sofer et al., 2015). Moreover, a cross-cultural study on Japanese and Israeli populations revealed that ingroup typical faces were perceived as more trustworthy than outgroup typical faces suggesting that people from different cultures use their own culture-specific typicality cues when judging trustworthiness (Sofer et al., 2017). Furthermore, facial distinctiveness and typicality have been repeatedly shown to be important for face recognition (Bartlett et al., 1984; Valentine, 1991; Valentine and Ferrara, 1991; Vokey and Read, 1992; O'toole et al., 1994; Burton et al., 2005; Valentine et al., 2016). Outgroup perception of typicality is suggested to be the core mechanism of racial stereotypes, where members of a minority that are perceived as more typical (of their own group) face a higher degree of racial prejudice and discrimination (Maddox, 2004; Kahn and Davies, 2011; Hebl et al., 2012). Nevertheless, our work does not focus on how typicality/distinctiveness affects the recognition of individual faces or on the stereotypicality of local minorities. Here we ask to what extent an individual face resembles the standards of ingroup and outgroup population. Our proposed approach is

not a refinement of intra-cultural facial typicality/distinctiveness research; rather, it is an extension of it into the area of cross-cultural comparisons.

In this research, we compare faces from two populations, Czech and Turkish that are not closely related but also not extremely distant as to the geographical distance as well as to the distance in similarity space. See **Figure 1** for illustration of differences between Czech and Turkish facial morphology. An individual's face may resemble the standards of facial appearance typical of a foreign population while at the same time being perceived as somewhat distinct within its own population. Therefore, we do not ask how distinct the face is from the facial average of its own population. Rather, the question is how to measure the deviances from morphological standards of own population *toward* the standards of some foreign population. This perspective is crucial if we want to catch the local dynamics of cross-cultural social perception. When a visitor arrives to a foreign country and is encountered by locals, his/her face is not compared to the standards of its home culture but to standards of the local culture, i.e., how his/her face is distinct from the local majority type. In our case, this corresponds to how much a Czech face looks Turkish-like and vice versa.

In general, there are four theoretical options for assessing typicality within a face space: (1) to measure Procrustes distance of all faces from global mean (average of all faces from both cultures under comparison); (2) to calculate distance from each face to local mean of its own population; (3) to calculate distance of each face from mean of outgroup (foreign) population (4) to project individual faces on the axis connecting the local means. Moreover, the objective measures of facial similarity based on Euclidean distances in principal component space has been shown to correspond to the human perception of facial similarity (Tredoux, 2002).

Options 1 and 2 do not allow to determine which of Czech faces are the most similar or dissimilar to Turkish standards and vice versa. The first option reveals only the similarity to a facial average which is the combination of intermediate Czech and Turkish features. The second option only informs about the similarity to local standards but tells nothing about the closeness of face to outgroup mean. For instance, a Czech face that is the most similar to Turks can have the same distance from its local average as the Czech face which is the most dissimilar, because faces are distributed radially to all directions around their local means. The distance from local average remains the best measure of within-culture typicality but does not allow cross-cultural comparison. The third and fourth options are more suitable for cross-cultural comparison because they provide some information about similarity of each face to the outgroup culture. The mean of one population is thus used as a reference to assess the level of distinctiveness of faces from second population and vice versa.

Inspired by research on human sexual dimorphism, we employed computational strategies represented by the third and fourth options that were formerly used for measuring the individual degree of male sex-typicality (Valenzano et al., 2006; Sanchez-Pages and Turiegano, 2010; Komori et al., 2011; Sánchez-Pagés and Turiegano, 2013; Sanchez-Pages et al., 2014;

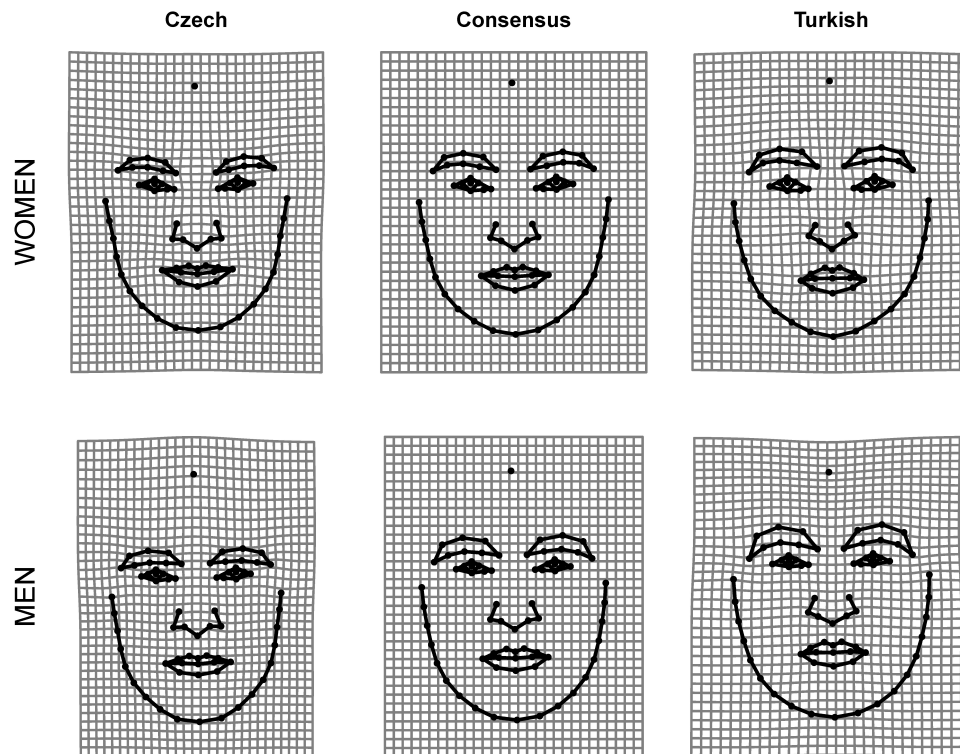


FIGURE 1 | Deformation grids illustrating the differences between Czech and Turkish mean facial shape compared to the global mean configuration. The shape changes were magnified three times for better interpretability.

Mitteroecker et al., 2015). Sanchez-Pages et al. (2014) suggested the possibility to use the distance between symmetrized facial configurations of each male individual and average female face as an objective measure of masculinity. In contrast, Mitteroecker et al. (2015) suggested the position of individual male faces along the axis between male and female average shape (maleness shape scores) as a measure of the individual degree of sexual dimorphism. This method, i.e., using two group averages to define an axis of morphological differences were formerly applied also by Komori et al. (2011) and Valenzano et al. (2006). The first approach is computationally identical to the option 3, i.e., distinctiveness measured as Distance from Outgroup Mean (DfOM) while the latter is identical to, option 4, a measure we call here as Cross-Group Typicality/Distinctiveness Metric (CTDM).

The aim of the present research is twofold: Study 1 aims to compare the measure of facial distinctiveness/typicality based on position of an individual face along the vector between ingroup and outgroup mean (CTDM) with a measure based on individual distance from the outgroup mean (DfOM). As the criterion measure, we gathered ratings of how foreign/local various ingroup and outgroup faces look in two cultures. We expect that CTDM will be more tightly correlated with these ratings than DfOM because DfOM does not carry any directional information about face position in morphospace. At the same time, we expect that the difference between correlations (CTDM – typicality/distinctiveness ratings vs.

DfOM – typicality/distinctiveness ratings) will be statistically significant.

In Study 2, we used manipulated composites based on different levels of CTDM to estimate the accuracy with which participants categorize outgroup vs. ingroup faces when they are paired. We expect that observers will generally be able to recognize the face of outgroups with accuracy higher than chance. Further, we expect that accuracy in this task will be lower for composite facial pairs showing lower CTDM distance (i.e., when both faces in the pair are closer to their respective outgroup means).

In sum, the main goal of this article is to propose a simple measure of distinctiveness and typicality which could be easily computed and, thanks to its one-dimensional (univariate) continuous nature, used as universal input variable, which reflects the individual degree of typicality/distinctiveness in any cross-group comparison, within all kinds of subsequent statistical modeling. In two studies, we aim to provide preliminary evidence that our new proposed measure behaves in line with theoretical expectations.

STUDY 1

Materials and Methods

In both studies, all procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration. The research

(ref. number 06/2017) was approved by The Institutional Review Board of Charles University Faculty of Science. The datasets generated for this study can be found in OSF¹.

Acquisition of Portrait Photographs

We obtained standardized portrait photographs of 100 men (50 Czech, Mean Age \pm SD = 23.89 ± 4.0 ; and 50 Turkish, Mean Age \pm SD = 21.54 ± 1.93) and 100 women (50 Czech, Mean Age \pm SD = 23.77 ± 4.32 ; and 50 Turkish, Mean Age \pm SD = 21.28 ± 1.34). The participants were instructed to avoid any facial cosmetics and jewelry, seated in front of a white background and asked to pose for the camera with neutral facial expression. The photographs of Czech Targets were taken with a Canon 6D camera using a 85 mm lens, studio flash, and a reflection screen. The distance from the lens to the face of the participant was 1.5 meters. A similar setup was used for collecting photographs of the Turkish targets. For the majority of photographs, a Nikon D90 with a 105 mm lens was used and targets were seated 2.82 meters from the camera (for details see Saribay et al., 2018). The photographs were cropped so that the eyes were in the same absolute height and the same length of neck was visible. The original image files had dimensions that were too large for subsequent online ratings. Therefore, the final image resolution was reduced to 600*745 pixels (width*height).

Portrait Ratings

Participants were sent an email inviting them to participate in an online study. A group of 315 Turkish raters (134 men, Mean Age \pm SD = 21.13 ± 2.09 and 181 women, Mean Age \pm SD = 20.9 ± 1.55) and 123 Czech raters (45 men, Mean Age \pm SD = 28.88 ± 12.83 and 78 women, Mean Age \pm SD = 27.45 ± 12.56) agreed to participate. Turkish raters were undergraduate students who participated in return for course credit and Czech raters were volunteers. Each rater was shown a total of 100 faces which were a random subset of 100 male and 100 female faces (50 of Turkish and Czech within each gender) in a randomized order, one face at a time. Raters were asked whether each face “looks more like a foreigner or more like a local person?” using a five-point response scale ranging from 1 (*certainly a foreigner*) to 5 (*certainly local*). No other information regarding the stimuli were given. Because higher ratings indicate greater certainty that the rated face belongs to the cultural ingroup, we refer to these ratings as “typicality/distinctiveness” and more specifically as “Turkishness” and “Czechness” when the assessment was done by Turkish and Czech raters, respectively. Inter-rater agreement estimated by Cronbach’s alpha was 0.968 for Turkishness and 0.882 for Czechness. Male and female ratings were highly correlated for both Czechness ($r = 0.926$, $p < 0.001$) and Turkishness ($r = 0.987$, $p < 0.001$).

Landmark Digitization and Procrustes Fit

Using TpsDig2 software (Rohlf, 2015), we defined 72 landmarks on each face. To specify the shape information along the curves 36 of total number of 72 landmarks were denoted semilandmarks. For definitions and positions of landmarks, see

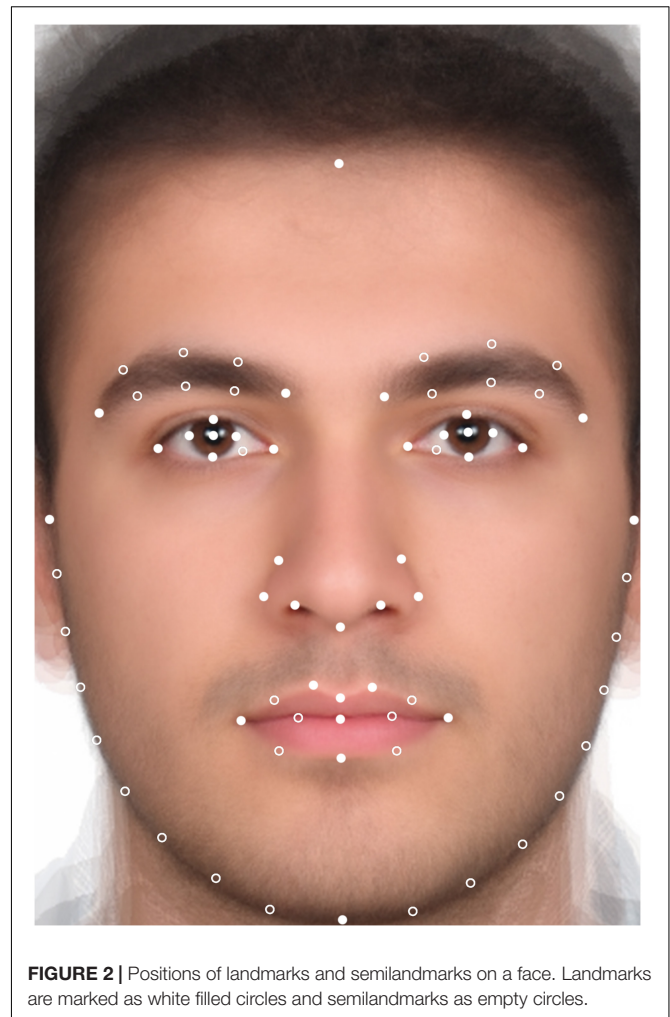


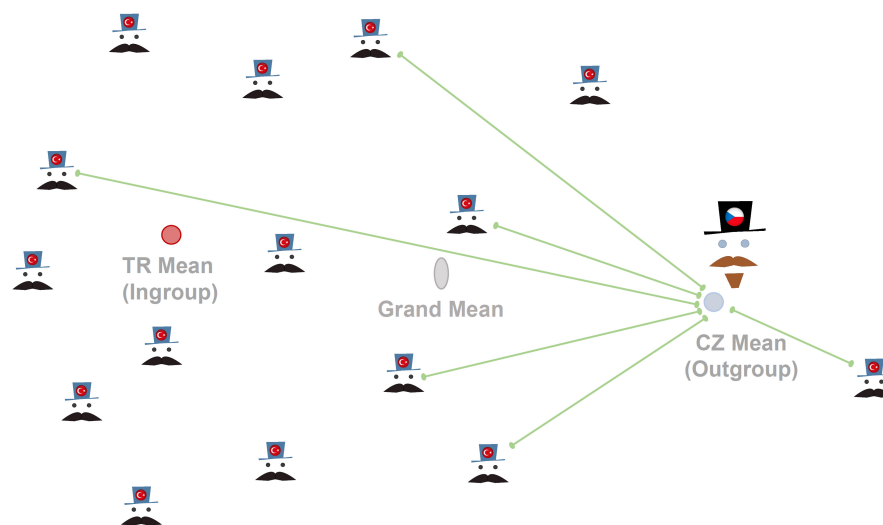
FIGURE 2 | Positions of landmarks and semilandmarks on a face. Landmarks are marked as white filled circles and semilandmarks as empty circles.

Figure 2 (the same configuration we used in our previous works, e.g., Kleisner et al., 2010; Danel et al., 2016).

All shape data were symmetrized and subsequently subjected to Generalized Procrustes Analysis using the ‘gpagen’ function implemented in the geomorph package in R (Adams and Otárola-Castillo, 2013). We pooled shape coordinates for Czech and Turkish facial configurations and ran GPA analysis on this joined dataset. GPAs were run separately for men and women to avoid shape variation due to sexual dimorphism. This procedure centered, scaled, and rotated all landmark configurations giving aligned shape coordinates (Procrustes residuals). The method that minimizes bending energy between each specimen and the Procrustes mean shape was used to slide the semilandmarks along the curves (Bookstein, 1997). The Procrustes-aligned data were projected to tangent space and used in subsequent multivariate analyses. For purposes of intrasexual comparisons of the two alternative measures of typicality/distinctiveness (DfOM vs. CTDM) the mean Czech and Turkish facial configuration (consensus) was computed separately for male and female faces. The average shape differences between Czech and Turkish facial configuration were visualized using thin plate spline (TPS)

¹<https://osf.io/wh7mf/>

A Distance from Outgroup Mean - DfOM



B Cross-group Typicality/Distinctiveness Metric – CTDM

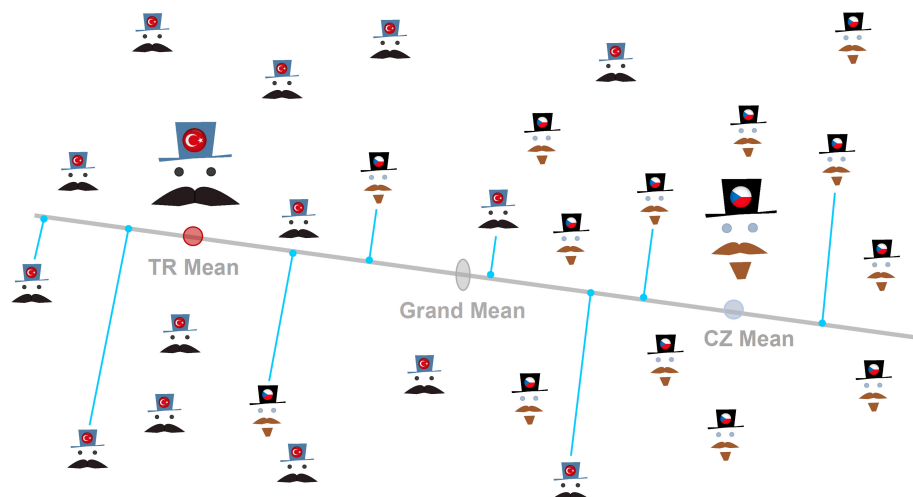


FIGURE 3 | Schematic representation of two approaches to measurement of distinctiveness/typicality compared in Study 1. Upper figure **(A)** demonstrates the approach (DfOM) where the distances of each Turkish face from outgroup mean (Czech average), i.e., green connecting lines, are used as a measure of individual distinctiveness. The shorter the distances (green line) the more distinct the face. Lower depiction **(B)** visualizes the suggested alternative approach (CTDM) taking the orthogonal projection (represented by blue lines) of Czech and Turkish faces on the mean difference vector as a measure of cross-culture distinctiveness. The projection of each face on the mean difference vector can be expressed by scores with negative values in direction to Turkish mean from grand mean and positive values toward the Czech mean. Note that in actual analyses the distances (green lines) or projections (blue lines) of all faces are calculated.

deformation grids (Bookstein, 1989; Rohlf and Marcus, 1993; Klingenberg, 2013).

Distance From Outgroup Mean (DfOM)

Distance from outgroup mean (DfOM) was computed as the Procrustes distance between the outgroup average facial

configuration (consensus) and each face in the set (**Figure 3A**). The outgroup defined relative to foreign faces at same time represents the ingroup for native faces. Thus the outgroup is understood from the perspective of the visitor to foreign country.

If Turkish faces are compared relative to Czech standards the DfOM is calculated as the distance of all faces

(Turkish and Czech) from Czech average and represents a measure that can be compared with ratings of “Czechness.” If Czech faces are compared relative to Turkish standards the DfOM is calculated as the distance of all faces (Turkish and Czech) from Turkish average and represents a measure that can be compared with ratings of “Turkishness.” The shorter is the distance of a face from outgroup consensus the more distinct is the face from its own population and the more typical it is of the foreign population.

Cross-Group Typicality/Distinctiveness Metric (CTDM)

To determine the position of an individual facial shape along the axis between ingroup and outgroup mean faces we projected the individual facial configurations in facial morphospace onto the vector between ingroup-outgroup means (see **Figure 3B**; see also Valenzano et al., 2006; Komori et al., 2011; Mitteroecker et al., 2015). This ‘between-group PCA’ represents a safer alternative to the linear discriminant analysis in cases where the number of individuals do not exceed the number of variables by several times (Mitteroecker and Bookstein, 2011). The position of an individual’s face along the axis connecting Czech and Turkish mean shape corresponds to relative distance of each facial configuration from average Czech and Turkish facial shape. This position can be numerically expressed by scores that correspond to projection of each face onto the principal components of the group averages. These scores vary with changes in facial morphology along the vector intersecting Czech and Turkish means and thus represent a one-dimensional proxy to overall multidimensional facial morphology. Higher negative scores indicate more Czech-like morphology whereas higher positive scores indicate more Turkish-like facial shape.

Statistical Analysis

All statistical procedures were performed in R 3.5.0. Kendall rank correlation coefficient was computed to measure the strength of relationship between variables. We compared Kendall’s correlations of CTDM and DfOM with ratings of Turkishness/Czechness (shared variable). The significance of the difference between the correlation coefficients together with their confidence intervals was bootstrapped. 10,000 random populations were sampled from the original data and the expected distribution of the coefficients and the difference between them was calculated. For the purpose of CI estimation, the bonds between ratings, CTDM, and DfOM values within individuals were maintained while sampling the individuals with replacement. This procedure was equivalent to the CI estimation in bootstrap version of “kendall.ci” function in NSM3 package in R (Schneider et al., 2018), which, however, does not take the relationship between more than two variables into consideration. When the distribution of expected differences between correlation coefficient was calculated, CTDM and DfOM vectors stayed unchanged to maintain their correlation within individuals and the ratings of Turkishness/Czechness was sampled with replacement.

To test for the shape differences between Czech and Turkish faces, we employed a permutation test based on comparison of the distance between Czech and Turkish means to the distances obtained by random assignment of observations to these groups. This was done, separately for each gender, by the “permudist” function implemented within the Morpho package in R (Schlager, 2017).

The differences in morphological variation between compared populations may have some effect on cross-cultural discrimination of faces because target faces from less variable population might have been easier to identify as belonging to the ingroup. Therefore, we tested for differences in morphological disparity between Czech and Turkish facial configurations using the function “morphol.disparity” in R’s geomorph package (Adams and Otárola-Castillo, 2013), with significance test based on 9,999 permutations. Using the same routine we also compared the morphological disparity of male and female faces because the facial traits responsible for populational identity might be easier to detect in the less variable gender.

Results and Discussion

Comparison of CTDM and DfOM

As expected, CTDM showed systematically tighter correlations with typicality/distinctiveness ratings (Czechness/Turkishness) than DfOM. The results of Kendall rank correlations (with bootstrapped CIs) are summarized in **Table 1**. See also **Supplementary Table 1** comparing the results of cross-group metrics (CTDM and DfOM) with distance calculated from ingroup means (DfIM). The bootstrap test of difference between Kendall’s correlations revealed that the association of typicality/distinctiveness ratings with CTDM is significantly stronger than with DfOM; this was true for faces of both men ($p = 0.013$) and women ($p = 0.029$).

TABLE 1 | Kendall’s rank correlations (with bootstrapped CIs) between Cross-Group Typicality/Distinctiveness Metric (CTDM), Distance from Outgroup Mean (DfOM), and ratings of typicality/distinctiveness by Turkish (Turkishness) and Czech (Czechness) raters.

	Male faces			
	Turkishness		Czechness	
	Kendall’s τ	CIs: 2.5% 97.5%	Kendall’s τ	CIs: 2.5% 97.5%
DfOM	0.196*	0.067 0.321	0.017	−0.113 0.147
CTDM	0.384**	0.271 0.488	−0.223*	−0.327 −0.117
	Female faces			
	Turkishness		Czechness	
	Kendall’s τ	CIs: 2.5% 97.5%	Kendall’s τ	CIs: 2.5% 97.5%
DfOM	0.258**	0.121 0.39	0.111	−0.018 0.242
CTDM	0.417**	0.331 0.496	−0.366**	−0.473 −0.254

See Methods for details on calculation of CTDM and DfOM. Note that CTDM may acquire negative values; higher negative scores indicate more Czech-like morphology whereas higher positive scores indicate more Turkish-like facial shape. * $p < 0.01$ and ** $p < 0.001$.

Morphological Differences Between Czech and Turkish Faces

Two group permutation test showed that Czech faces significantly differ in their facial shape from Turkish faces both for faces of men (Procrustes distance between means–PDM = 0.01995, $p < 0.001$) and women (PDM = 0.01634, $p < 0.001$). This means that the members of our target populations differ in average as to their facial structure which makes the cross-group comparison morphologically meaningful.

The analysis of morphological disparity (MD) showed that Czech men (MD = 0.00199) are more variable in facial shape than Turkish men (MD = 0.00154) and Czech women (MD = 0.00182) are more variable in facial shape than Turkish women (MD = 0.00142). These differences were significant for both male ($p = 0.015$) and female ($p = 0.01$) faces. When sex differences were tested, men and women (including both Czech and Turkish faces) did not differ significantly as to the variation in facial shape ($p = 0.149$). These results suggest that Czechs are generally more variable in facial morphology than Turks. The raters might be thus more effective in classifying the Turkish faces as they are (at least in our sample of faces) morphologically more homogeneous. The stronger effects reported for Turkish faces seems to support this expectation.

STUDY 2

Materials and Methods

Production of Manipulated Stimuli

We generated manipulated composites in order to estimate the accuracy with which participants categorize a face across different types of trials where the following types of Turkish (TR) and Czech (CZ) face composites are paired: 1: CZ farthest away from TR; 2: Average CZ; 3: CZ closest to TR; 4: TR closest to CZ; 5: Average TR; 6: TR farthest away from CZ (when different types of faces are paired; e.g., 1 vs. 4).

The composite faces were generated with use of TpsSuper 2.05 software (Rohlf, 2015). Six facial images of individuals closest to a selected position determined by CTDM value (i.e., CZ farthest away from TR mean; Average CZ; CZ closest to TR mean; TR closest to CZ mean; Average TR; TR farthest away from CZ mean) were used to create composites. See **Figure 4** for exposition of manipulated composites and **Supplementary Figures 1, 2**, men and women, respectively, for the position of composites along CTDM axis.

Two-Alternative Forced-Choice (2AFC) Discrimination of Manipulated Stimuli

Participants were sent an email inviting them to participate in an online study. 327 Turkish university undergraduates (153 men, Mean Age \pm SD = 20.63 \pm 1.46; 171 women, Mean Age \pm SD = 20.6 \pm 1.8; 3 others, Mean Age \pm SD = 23 \pm 3.6) participated in return for course credit. Participants were asked to view composite face pairs (whose typicality/distinctiveness was manipulated based on CTDM) and to select the face in each pair that “looks more foreign.” We focused only on cross-group pairs. That is, in each pair, one face belonged objectively to the CZ

group and the other to the TR group (with left-right position on the computer screen counterbalanced). This resulted in the following combinations (in terms of the above categories; see *Production of manipulated stimuli* section and **Figure 4**): 1 vs. 4; 1 vs. 5; 1 vs. 6; 2 vs. 4; 2 vs. 5; 2 vs. 6; 3 vs. 4; 3 vs. 5; 3 vs. 6. One trial for each possible combination was shown to each participant, in random order. Accuracy was defined in terms of whether the face selected as looking more foreign is objectively an outgroup (i.e., Czech) face.

Statistical Analysis

All statistical procedures were performed in R 3.5.0. The average difference in proportion of correct responses in trials in which the Czech face was presented on the left vs. right of the screen was -0.0157 (range = -0.059 to 0.036), suggesting that location of faces did not have a systematic effect on responses. Thus, we analyzed the data collapsing across this counterbalancing factor.

Overall strategy

We were interested in (1) whether participants were generally able to accurately distinguish outgroup from ingroup faces and (2) whether there would be any difference in the level of this ability in trials with different face pairs. We checked the former using one-proportion z-tests against chance level performance (i.e., 50% accuracy). For the latter, we first conducted Cochran's Q tests (with the RVAideMemoire package in R; Hervé, 2018) to examine whether any two face pairs differed in the proportion of correct responses that participants produced in those trials. To follow-up the significant Cochran's Q tests, we conducted a set of selective pairwise comparisons. For each gender (of face), 9 pairs had been presented to participants, resulting in 36 possible pairwise comparisons in total, most of which are theoretically uninteresting or uninterpretable. For ease of interpretation, we focused on comparisons in which we could hold one face constant (e.g., 1_4 vs. 3_4). We selected specific pairs of face pairs to explore the idea that Turkish raters would have more difficulty (i.e., produce lower proportion of correct responses) in trials that involved the Czech face closest to the Turkish average (#3), compared to the Czech faces that are further away from the Turkish average (#1 and #2). That is, the Czech face closest to the Turkish average should be more often mistakenly judged as not being foreign, compared to Czech faces further away from average, holding constant the Turkish face in the pair. We would expect this to be especially true compared to the Czech face furthest away from the Turkish average (#1) and to a lesser extent, to the average Czech face (#2).

This highlights the following 6 comparisons (per gender of face), with the expectation that there would be lower proportion of correct responses in the second pair in each comparison: 1_X vs. 3_X and 2_X vs. 3_X; where X takes the values of 4, 5, and 6 in different trials. We conducted these pairwise comparisons using McNemar's test (asymptotic; using the rcompanion R package; Mangiafico, 2018), which tests the null hypothesis that the proportion of responses is equal across two face pairs compared. To control for familywise error rate, p -values were adjusted using the Hochberg method. We also employed binomial mixed effects models (“glmer” function within lmerTest package;



FIGURE 4 | Visualization of manipulated composite stimuli of men (above) and women (below) based on different levels of CTDM along mean difference vector.

Kuznetsova et al., 2017) with “responses to CTDM trials” (coded as 1 = true answer, 0 = false answer) as a response variable, “type of CTDM trials” as an independent variable, and “raters identity” as a random effect. We have built separate models for male and female facial stimuli. We specified relevant contrasts between trials using “glht” function within multcomp package (Hothorn et al., 2008).

Data exclusions

Because the findings could be affected by the presence of raters whose cultural ingroup is other than Turkish, we sought to exclude such raters from the dataset. Participants were asked to report their nationality but this did not capture cultural ingroup-outgroup status or possible differences in amount of exposure to Turkish individuals (e.g., some participants who indicated having a nationality other than Turkish were dual citizens). Instead, we calculated approximately how many months of their lifetime each participant had spent in their homeland by taking the difference between their age in months and the months that they reported being abroad. There were 6 participants who had spent a month or less in Turkey. We excluded these participants from further analyses because their exposure to the Turkish faces appears to be very limited based on this information (1 participant with missing data on this measure was retained).

In the remaining sample, average proportion of correct responses for each participant across the 18 trials of the 2AFC task ranged from 0.16 to 1. There were only 6 participants with an average proportion of correct responses lower than 0.5. The morphological closeness of face pairs presented in the 2AFC trials and hence the difficulty of accurately responding suggests that these low-performers were not necessarily careless or responding randomly. Thus, we retained these low-performing participants. The final sample consisted of 321 participants. These decisions were inconsequential as the findings were not affected

by inclusion vs. exclusion of participants on the basis of limited exposure to Turkish faces or low performance in the 2AFC (e.g., the decisions to reject or retain the null hypothesis were unchanged in Table 3).

Results and Discussion

Ability to Distinguish Foreign Faces

We checked whether participants were generally able to distinguish foreign faces from local ones. Table 2 presents the percentage of correct responses in the 2AFC trials across different face pairs separately for male and female target faces. Overall, participants appeared to have no difficulty distinguishing foreign faces. One-proportion z-tests indicated that the proportions for all face pairs in Table 2 were significantly greater than chance level (0.5) ($p_s < 0.0001$).

TABLE 2 | Proportion of correct responses in the 2AFC for each face pair by face gender.

Face pair	Male faces	Female faces
1_4	0.95	0.944
1_5	0.96	0.953
1_6	0.947	0.857
2_4	0.91	0.96
2_5	0.894	0.966
2_6	0.866	0.854
3_4	0.903	0.785
3_5	0.875	0.866
3_6	0.897	0.701
Average within Gender of face	0.911	0.876
Global average	0.894	

See Figure 4 or the text for the correspondence between values in the face pair column and composite images.

TABLE 3 | Pairwise comparisons (McNemar's Tests) of correct response proportions in trials with different face pairs.

Comparison pair	Male faces		Female faces	
	χ^2	Adjusted p	χ^2	Adjusted p
1_4 vs. 3_4	6.43	0.045*	37.7	< 0.001***
2_4 vs. 3_4	0.1	0.752	44.8	< 0.001***
1_5 vs. 3_5	18.69	0.000***	18.67	< 0.0010***
2_5 vs. 3_5	0.783	0.752	22.26	< 0.001***
1_6 vs. 3_6	9.14	0.013*	33.78	< 0.001***
2_6 vs. 3_6	3.33	0.204	36.94	< 0.001***

Hochberg-adjusted p-values: * $p < 0.05$; ** $p < 0.01$; and *** $p < 0.001$.

Differences in Ability to Distinguish Foreign Faces Across Face Pairs

Next, we examined whether participant's ability to produce the correct response (i.e., choosing the outgroup face as "foreign") differed across face pairs. For this purpose, we conducted a Cochran's Q test as an omnibus test of whether the proportion of correct responses for any pairs were different within the 9 pairs of male faces and 9 pairs of female faces, separately. The proportion of correct responses differed between at least one comparison within both male, $\chi^2(8, N = 321) = 52.2, p < 0.0001$; and female face pairs, $\chi^2(8, N = 321) = 235, p < 0.0001$.

To follow up, we compared selected pairs of face pairs (see *Statistical Analysis* section above) using McNemar's test. These pairwise comparisons are listed in **Table 3**. Together with the proportions listed in **Table 2**, it can be seen that the proportion of correct responses in all significant comparisons are in the expected direction. The comparisons that we expected to be most different are all significant for male faces; whereas the comparisons for which we had less strong expectation of difference are not. For female faces, the comparisons for all pairs were significantly different. The result carried out by binomial mixed effect modeling corroborated the results of the McNemar's test (see **Supplementary Tables 2, 3**).

There could be limitations arising from the design that could be addressed in future research for a better, confirmatory test of the idea. In sum, while there are apparently other sources of variance in responses in the 2AFC, there is at least preliminary evidence from these results that outgroup faces closer to the ingroup average are more difficult to correctly categorize as "foreign" and/or that outgroup faces further away from the ingroup average are easier to correctly categorize as "foreign."

GENERAL DISCUSSION

In the present research, we took advantage of geometric morphometrics to propose a sample dependent and thus data-driven method for estimating the individual degree of facial distinctiveness/typicality for cross-cultural comparisons. We attempted to provide preliminary evidence in support of a novel cross-cultural metric of typicality/distinctiveness (CTDM) in two studies. Study 1 revealed significantly tighter association between ratings of typicality/distinctiveness and CTDM than the same

ratings and an alternative approach (DfOM). The same pattern was found for faces of both men and women and for both Turkish and Czech raters. The possible weakness of DfOM is that a pure distance measure does not carry information about the position of a face in facial morphospace. The reason why CTDM showed stronger correlations with ratings than DfOM is that DfOM is based on Procrustes distance and can have only positive values. That is, it provides no (or very limited) information about mutual positions of targets (faces) and reference (outgroup mean). The straight lines between individual faces and outgroup mean may form variety of angles with between group axis including right angle. DfOM thus does not account for the fact that faces may theoretically vary in all directions from outgroup mean (see **Figure 3A**). For instance, two Turkish faces having the same distance from the Czech mean may look dissimilar because one of the faces could be even less similar to Turkish standards than the Czech mean. In contrast, CTDM aligns the positional information of individual faces along the cross-group axis (see **Figure 3B**). The above-mentioned problematic interpretation of two Turkish faces having the same DfOM despite their apparent dissimilarity thus becomes clearer within the CTDM framework wherein these two faces acquire different CTDM scores.

In Study 2, we assessed accuracy of discrimination of outgroup vs. ingroup composite faces varying along CTDM in a 2AFC task. Performance was generally consistent with our expectations. For instance, when pairs of face pairs were compared, the pairs that contrasted faces closer in CTDM distance (e.g., the Czech face closest to Turkish average and the Turkish face closest to Czech average) were harder to accurately discriminate compared to face pairs in which the two faces are further apart from each other in CTDM. Results involving female (vs. male) faces were more consistent with expectations. This may indicate that female faces carry more information about cultural identity, a possibility that future research could examine.

In hindsight, even though participants showed high overall accuracy, the 2AFC task may have been difficult for participants as the Czech and Turkish faces show only slightly difference in morphology (see **Figure 4**). Other features of task design may have not been optimal for the current purposes, as well. For instance, because the same faces were repeatedly used across trials, participants may have anchored their responses for repetitions of a given face on their response in the first trial with that face (i.e., consistently judging face 1 to be foreign across trials where it was paired with 4, 5, and 6). This could have masked variability in responses that would otherwise occur. Future tests could better control such irrelevant task features for a purer test of CTDM effects. Despite these limitations, we view Study 2 findings as providing encouraging preliminary evidence for the usefulness of CTDM in predicting performance in ingroup-outgroup face discrimination situations.

Future Directions, Limitations, and Potential Applications

We used only the shape information for calculation of CTDM but faces are more complex. Hence, the face space can theoretically be augmented with further types of non-shape variables and CTDM may be computed based on this improved face space. Future

studies should take into account also facial size, texture, skin color, eye color, hair style and color, contrasts between mouth and surrounding skin, and so on.

Cross-Group Typicality/Distinctiveness Metric can have a broad range of social applications. Previous research showed positive association between judgments of trustworthiness and facial typicality (Sofer et al., 2015, 2017). In cross-cultural interactions, such as bilateral business tradeoffs, student exchange, or even war conflicts, the outgroup individuals closest to the ingroup standards would obtain a substantial advantage because their faces will be more trustworthy-looking to potential, business partners, tutors, or invaders, respectively. In fact, a given face being typical of the ingroup should engender feelings of familiarity, which is known to underlie positive responses to that face (Zebrowitz et al., 2008). Viewed differently, a lack of such typicality and familiarity for an encountered face and/or the face resembling an outgroup prototype is known to make prejudiced reactions toward the face-bearer more likely, sometimes even rendering the situation into a matter of life and death (Blair et al., 2004; Eberhardt et al., 2006). Importantly, these typicality/familiarity-contingent responses to faces may go beyond mere categorization of faces as ingroup vs. outgroup: Perceivers' responses are often driven by facial cues, not necessarily by the perceived social category of faces (Livingston and Brewer, 2002). The influence of group typicality on social perception had long been ignored with the field of social psychology but has been discussed more explicitly in recent years (Maddox, 2004). However, typicality is often coded in a categorical fashion (i.e., the main distinction being between individuals who are more vs. less typical of a group; e.g., Hebl et al., 2012; Study 2) and/or relying on human judges (e.g., Hebl et al., 2012; Studies 1 and 3). Furthermore, such studies usually focus on only one side of the coin: How typical of the ingroup individuals appear. In the case of research on prejudice against African-Americans in the United States, researchers may implicitly have in mind "smaller distance from White American average" when they mention "low racial stereotypicality" of African-American faces, but this is not made explicit (an African-American face can also be less typical of the ingroup and be closer to a different ethnic group such as Hispanic). CTDM may offer a way to distinguish these "cue- vs. category-based" responses and the influence of individual differences in targets' appearance in a way that is conceptually clearer, more quantitative and/or objective, and more finely (vs. coarsely) related to face morphology than previous research. A Turkish visitor in Czech Republic should thus receive better treatment from Czech individuals when his/her CTDM is closer to (vs. farther away from) the Czech populational average, in a way that goes beyond him/her being recognized as Turkish (or foreign). In short, CTDM could be used to predict prejudice in intergroup contexts beyond the effect of categorization and could have wide application in social psychology.

Another possible application may link CTDM with face recognition. According to the attractor field model, the object's similarity and spatial density in multidimensional space is unintuitively interrelated. Face morphs were judged to be more similar to the atypical than to the typical parent image (Bartlett and Tanaka, 1998; Tanaka et al., 1998). It is thus

more difficult to detect differences between atypical morphs than differences between typical morphs. This was true for morphs of various inanimate and animate objects such as birds, cars, and faces (Tanaka and Corneille, 2007). There seems to be a tradeoff between spatial density (and higher similarity) of faces around local means and bigger attractor fields around distinctive faces. These distinctive faces thus may be perceived as mutually more similar due to their larger attractor fields than they 'objectively' are. Thereof stems one prediction for future research testable with use of CTDM: The individual identity of outgroup faces closest to observers' ingroup mean should be recognized with higher accuracy than individual identities of faces most distant from observers' ingroup standards. A Turkish visitor in Czech Republic should thus better recognize the identity of Czech individuals with CTDM values closer to Turkish populational average.

As our research was intended as an initial test of CTDM, there were limitations other than those already discussed. Some differences between images from the two cultural groups need to be better controlled. Even minor differences in technical equipment used to produce images (e.g., focal length of the camera lens) could result in different images (Třebický et al., 2016). Other stylistic differences such as facial hair may also influence ratings. Future research should seek to remedy these problems.

Further Theoretical Considerations: A Note on the Limitations of Composite Images in Face Research

Portrait photograph blending is an old procedure, first used by Francis Galton who applied it to reveal features typical of certain categories of people (Galton, 1879, 1883). A long time since Galton, average composites are still used, with much technical improvement. Many researchers use manipulated composite stimuli to investigate various causative effects of facial traits on social impressions. We see this research agenda as at least partially problematic (see also Schaefer et al., 2009; Jones, 2018; Windhager et al., 2018). Calibrating facial morphs for use as stimuli in biological studies of social perception. Scientific reports, 8(1), 6698. doi: 10.1038/s41598-018-24911-0. Unlike individual images, composites yield clear results even when the sample size is low. For instance, Rhodes (2006) reported a strong positive relationship between symmetry and attractiveness in composite faces, but only moderate effects in non-manipulated faces (Rhodes, 2006). Moreover, recent cross-cultural evidence showed only moderate or no relationship between attractiveness ratings based on non-manipulated facial photos and averageness computed as each face's distance from sample Procrustes mean (Kleisner et al., 2017). The use of experimentally manipulated stimuli thus has various practical consequences, such as greater effect sizes and a higher probability of positive results. Moreover, stimuli experimentally manipulated for a particular research purpose become a reification of researchers' theoretical needs. This is a problem when such stimuli are not used as research tools but substituted for natural objects (faces), that is, when individual faces with their natural variations are substituted for manipulated

stimuli whose variation is constrained in a way that *a priori* corresponds to expectations given by a theory. By constraining the variation of stimuli, we also limit the variation of possible responses to these stimuli. One might claim that this is how experimental science works, which may well be so, but properties of the experimental toolkit must be included in the interpretation of the results. During twentieth century, this condition has been widely discussed within philosophy of science and become an indispensable part of some fields of experimental physics. Yet, it remains largely neglected in evolutionary psychology and biology.

What is the alternative? First, to use non-manipulated stimuli. Second, to use stimuli manipulated so as to correspond to the observed range of natural variation. Third, to use both manipulated and non-manipulated stimuli; the difference in results could then be used as a background to the overall interpretation of results. This can be easily accomplished by application of geometric morphometrics and related multivariate techniques, such as in the CTDM approach presented here, that provide direct statistical control over the direct stimuli analysis and their manipulation.

CONCLUSION

To conclude, distinctiveness and typicality are two sides of the same coin so claiming that a face is either more/less distinct or more/less typical depends on which populations are taken as ingroup and outgroup. CTDM allows one to estimate the degree to which an individual from a given (ingroup/local) population resembles the facial standards of another (outgroup/foreign) population and vice versa. When mathematically expressed, such knowledge is potentially useful for studying relationships between the individuals' degree of cultural distinctiveness/typicality perceived by others and attributions of attractiveness and personality traits across cultures. Further, CTDM allows us to generate manipulated stimuli that respect the natural variation of human faces within a particular population. Finally, CTDM is not constrained to human faces and can be applied to any shape such as parts of the human body and cultural artifacts. We hope that future research will provide further evidence of CTDM's utility and realize its potential for application in face research and beyond.

REFERENCES

- Adams, D. C., and Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* 4, 393–399. doi: 10.1111/2041-210X.12035
- Bartlett, J. C., Hurry, S., and Thorley, W. (1984). Typicality and familiarity of faces. *Mem. Cogn.* 12, 219–228. doi: 10.3758/BF03197669
- Bartlett, M. S., and Tanaka, J. W. (1998). "An attractor field model of face representation: effects of typicality and image morphing," in *Psychonomics Society Satellite Symposium on Object Perception and Memory (OPAM)*, Dallas, TX.
- Blair, I. V., Judd, C. M., and Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychol. Sci.* 15, 674–679. doi: 10.1111/j.0956-7976.2004.00739.x

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of "The Institutional Review Board of Charles University Faculty of Science" with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the "The Institutional Review Board of Charles University Faculty of Science" (protocol ref. number: 06/2017).

AUTHOR CONTRIBUTIONS

KK and SAS developed the study concept and drafted the manuscript. Data collection was performed by SAS in Turkey and by ŠP and KK in Czechia. All authors performed the data analysis and result interpretation, provided critical revisions, and approved the final version of the manuscript for submission.

FUNDING

This study was supported by Czech Science Foundation grant number GA18-10298S.

ACKNOWLEDGMENTS

We thank Petr Tureček for his development of the bootstrapped test of Kendall's tau differences used in Study 1, Salvatore Mangiafico for sending us a modified version of his rcompanion package that outputs exact test values for the McNemar's tests used in Study 2, and Jakub Kreisinger for helpful advices on binomial mixed effect modeling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00124/full#supplementary-material>

- Bookstein, F. L. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 567–585. doi: 10.1109/34.24792
- Bookstein, F. L. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Image Anal.* 1, 225–243. doi: 10.1016/S1361-8415(97)85012-8
- Burton, A. M., Jenkins, R., Hancock, P. J., and White, D. (2005). Robust representations for face recognition: the power of averages. *Cogn. Psychol.* 51, 256–284. doi: 10.1016/j.cogpsych.2005.06.003
- Danel, D., Dziedzic-Danel, A., and Kleisner, K. (2016). Does age difference really matter? Facial markers of biological quality and age difference between husband and wife. *HOMO J. Comp. Hum. Biol.* 67, 337–347. doi: 10.1016/j.jchb.2016.05.002

- Danel, D. P., Fedurek, P., Coetzee, V., Stephen, I. D., Nowak, N., Stirrat, M., et al. (2012). A cross-cultural comparison of population-specific face shape preferences (Homo sapiens). *Ethology* 118, 1173–1181. doi: 10.1111/eth.12022
- DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., and Feinberg, D. R. (2007). Dissociating averageness and attractiveness: attractive faces are not always average. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1420–1430. doi: 10.1037/0096-1523.33.6.1420
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., and Johnson, S. L. (2006). Looking deathworthy: perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychol. Sci.* 17, 383–386. doi: 10.1111/j.1467-9280.2006.01716.x
- Fink, B., and Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Curr. Dir. Psychol. Sci.* 11, 154–158. doi: 10.1111/1467-8721.00190
- Galton, F. (1879). Composite portraits, made by combining those of many different persons into a single resultant figure. *J. Anthropol. Inst. Great Br. Ireland* 8, 132–144. doi: 10.2307/2841021
- Galton, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan and Company. doi: 10.1037/14178-000
- Goethe, J. W. (1999). *Die Schriften zur Naturwissenschaft*. Stuttgart: Reclam.
- Hebl, M. R., Williams, M. J., Sundermann, J. M., Kell, H. J., and Davies, P. G. (2012). Selectively friending: racial stereotypicality and social rejection. *J. Exp. Soc. Psychol.* 48, 1329–1335. doi: 10.1016/j.jesp.2012.05.019
- Hervé, M. (2018). *RVAideMemoire: Testing and Plotting Procedures for Biostatistics*. Available at: <https://cran.r-project.org/web/packages/RVAideMemoire/>
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometric J.* 50, 346–363. doi: 10.1002/bimj.200810425
- Jones, A. L. (2018). Beyond average: using face regression to study social perception. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/dpmzq
- Kahn, K. B., and Davies, P. G. (2011). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Process. Intergroup Relat.* 14, 569–580. doi: 10.1177/1368430210374609
- Kleisner, K. (2007). The formation of the theory of homology in biological sciences. *Acta Biotheor.* 55, 317–340. doi: 10.1007/s10441-007-9023-8
- Kleisner, K., Kočnar, T., Rubesová, A., and Flegel, J. (2010). Eye color predicts but does not directly influence perceived dominance in men. *Pers. Individ. Differ.* 49, 59–64. doi: 10.1016/j.paid.2010.03.011
- Kleisner, K., Kočnar, T., Tureček, P., Stella, D., Akoko, R. M., Třebický, V., et al. (2017). African and European perception of African female attractiveness. *Evol. Hum. Behav.* 38, 744–755. doi: 10.1016/j.evolhumbehav.2017.07.002
- Klingenberg, C. P. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix* 24, 15–24.
- Komori, M., Kawamura, S., and Ishihara, S. (2011). Multiple mechanisms in the perception of face gender: effect of sex-irrelevant features. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 626–633. doi: 10.1037/A0020369
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13
- Langlois, J. H., and Roggman, L. A. (1990). Attractive faces are only average. *Psychol. Sci.* 1, 115–121. doi: 10.1111/j.1467-9280.1990.tb00079.x
- Livingston, R. W., and Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *J. Pers. Soc. Psychol.* 82, 5–18. doi: 10.1037/0022-3514.82.1.5
- Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Pers. Soc. Psychol. Rev.* 8, 383–401. doi: 10.1207/s15327957pspr0804_4
- Mangiafico, S. (2018). *Rcompanion: Functions to Support Extension Education Program Evaluation*. R Package Version 2.0.0. Available at: <https://cran.r-project.org/web/packages/rcompanion/>
- Mitteroecker, P., and Bookstein, F. (2011). Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. *Evol. Biol.* 38, 100–114. doi: 10.1007/s11692-011-9109-8
- Mitteroecker, P., Windhager, S., Müller, G. B., and Schaefer, K. (2015). The morphometrics of “masculinity” in human faces. *PLoS One* 10:e0118374. doi: 10.1371/journal.pone.0118374
- O’toole, A. J., Deffenbacher, K. A., Valentin, D., and Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Mem. Cogn.* 22, 208–224. doi: 10.3758/BF03208892
- Perrett, D. I., May, K. A., and Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature* 368, 239–242. doi: 10.1038/368239a0
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Rhodes, G., and Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychol. Sci.* 7, 105–110. doi: 10.1111/j.1467-9280.1996.tb00338.x
- Rohlf, F. J. (2015). The tps series of software. *Hystrix, the Italian J. Mammal.* 26, 9–12.
- Rohlf, J. F., and Marcus, L. F. (1993). A revolution morphometrics. *Trends Ecol. Evol.* 8, 129–132. doi: 10.1016/0169-5347(93)90024-J
- Russell, E. S. (1916). *Form and Function: A Contribution to the History of Animal Morphology*. London: Murray.
- Said, C. P., and Todorov, A. (2011). A statistical model of facial attractiveness. *Psychol. Sci.* 22, 1183–1190. doi: 10.1177/0956797611419169
- Sanchez-Page, S., Rodriguez-Ruiz, C., and Turiegano, E. (2014). Facial masculinity: how the choice of measurement method enables to detect its influence on behaviour. *PLoS One* 9:e112157. doi: 10.1371/journal.pone.0112157
- Sanchez-Page, S., and Turiegano, E. (2010). Testosterone, facial symmetry and cooperation in the prisoners’ dilemma. *Physiol. Behav.* 99, 355–361. doi: 10.1016/j.physbeh.2009.11.013
- Sánchez-Pagés, S., and Turiegano, E. (2013). Two studies on the interplay between social preferences and individual biological features. *Behaviour* 150, 713–735. doi: 10.1163/1568539X-00003077
- Saribay, S. A., Biten, A. F., Meral, E. O., Aldan, P., Třebický, V., and Kleisner, K. (2018). The Bogazici face database: standardized photographs of Turkish faces with supporting materials. *PLoS One* 13:e0192018. doi: 10.1371/journal.pone.0192018
- Schaefer, K., Mitteroecker, P., Fink, B., and Bookstein, F. L. (2009). Psychomorphospace—from biology to perception, and back: towards an integrated quantification of facial form variation. *Biol. Theory* 4, 98–106. doi: 10.1162/biot.2009.4.1.98
- Schlager, S. (2017). “Morpho and Rvcg – shape analysis in R,” in *Statistical Shape and Deformation Analysis: Methods, Implementation and Applications*, eds G. Zheng, S. Li, and G. Székely (Cambridge, MA: Academic Press), 217–256.
- Schneider, G., Chicken, E., and Becvarik, R. (2018). *NSM3: Functions and Datasets to Accompany Hollander, Wolfe, and Chicken – Nonparametric Statistical Methods, Third Edition*. R Package Version 1.12. Available at: <https://cran.r-project.org/web/packages/NSM3/>
- Sofer, C., Dotsch, R., Oikawa, M., Oikawa, H., Wigboldus, D. H., and Todorov, A. (2017). For your local eyes only: culture-specific face typicality influences perceptions of trustworthiness. *Perception* 46, 914–928. doi: 10.1177/0301006617691786
- Sofer, C., Dotsch, R., Wigboldus, D. H., and Todorov, A. (2015). What is typical is good: the influence of face typicality on perceived trustworthiness. *Psychol. Sci.* 26, 39–47. doi: 10.1177/0956797614554955
- Tanaka, J., Giles, M., Kremen, S., and Simon, V. (1998). Mapping attractor fields in face space: the atypicality bias in face recognition. *Cognition* 68, 199–220. doi: 10.1016/S0010-0277(98)00048-1
- Tanaka, J. W., and Corneille, O. (2007). Typicality effects in face and object perception: further evidence for the attractor field model. *Percept. Psychophys.* 69, 619–627. doi: 10.3758/BF03193919
- Třebický, V., Fialová, J., Kleisner, K., and Havlíček, J. (2016). Focal length affects depicted shape and perception of facial images. *PLoS One* 11:e0149313. doi: 10.1371/journal.pone.0149313
- Tredoux, C. (2002). A direct measure of facial similarity and its relation to human similarity perceptions. *J. Exp. Psychol. Appl.* 8, 180–193. doi: 10.1037/1076-898X.8.3.180
- Trujillo, L. T., Jankowitsch, J. M., and Langlois, J. H. (2014). Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cogn. Affect. Behav. Neurosci.* 14, 1061–1076. doi: 10.3758/s13415-013-0230-2

- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q. J. Exp. Psychol. Sec. A* 43, 161–204. doi: 10.1080/14640749108400966
- Valentine, T., and Ferrara, A. (1991). Typicality in categorization, recognition and identification: evidence from face recognition. *Br. J. Psychol.* 82, 87–102. doi: 10.1111/j.2044-8295.1991.tb02384.x
- Valentine, T., Lewis, M. B., and Hills, P. J. (2016). Face-space: a unifying concept in face recognition research. *Q. J. Exp. Psychol.* 69, 1996–2019. doi: 10.1080/17470218.2014.990392
- Valenzano, D. R., Mennucci, A., Tartarelli, G., and Cellerino, A. (2006). Shape analysis of female facial attractiveness. *Vis. Res.* 46, 1282–1291. doi: 10.1016/j.visres.2005.10.024
- Vokey, J. R., and Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Mem. Cogn.* 20, 291–302. doi: 10.3758/BF03199666
- Wickham, L. H., Morris, P. E., and Fritz, C. O. (2000). Facial distinctiveness: its measurement, distribution and influence on immediate and delayed recognition. *Br. J. Psychol.* 91, 99–123. doi: 10.1348/000712600161709
- Windhager, S., Bookstein, F. L., Mueller, H., Zunner, E., Kirchengast, S., and Schaefer, K. (2018). Calibrating facial morphs for use as stimuli in biological studies of social perception. *Sci. Rep.* 8:6698. doi: 10.1038/s41598-018-24911-0
- Zebrowitz, L. A., White, B., and Wieneke, K. (2008). Mere exposure and racial prejudice: exposure to other-race faces increases liking for strangers of that race. *Soc. Cognit.* 26, 259–275. doi: 10.1521/soco.2008.26.3.259

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Kleisner, Pokorný and Saribay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



She Always Steps in the Same River: Similarity Among Long-Term Partners in Their Demographic, Physical, and Personality Characteristics

Zuzana Štěrbová^{1,2*}, Petr Tureček^{2,3} and Karel Kleisner^{2,3}

¹ Department of Zoology, Faculty of Science, Charles University, Prague, Czechia, ² National Institute of Mental Health, Klecany, Czechia, ³ Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czechia

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Christopher Watkins,
Abertay University, United Kingdom
Tamsin Saxton,
Northumbria University,
United Kingdom

*Correspondence:

Zuzana Štěrbová
zuzana.sterbova@natur.cuni.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 October 2018

Accepted: 09 January 2019

Published: 05 February 2019

Citation:

Štěrbová Z, Tureček P and
Kleisner K (2019) She Always Steps
in the Same River: Similarity Among
Long-Term Partners in Their
Demographic, Physical,
and Personality Characteristics.
Front. Psychol. 10:52.
doi: 10.3389/fpsyg.2019.00052

In mate choice, individuals consider a wide pool of potential partners. It has been found that people have certain preferences, but intraindividual stability of mate choice over time remains little explored. We tested individual consistency of mate choice with respect to a number of demographic, physical, and personality characteristics. Only mothers were recruited for this study, because we wanted to find out not only whether women choose long-term partners with certain characteristics but also whether the father of their child(ren) differs from their other long-term (ex-)partners. Women ($N = 537$) of 19–45 years of age indicated the demographic, physical (by using image stimuli), and personality characteristics of all of their long-term partners (partners per respondent: mean = 2.98, $SD = 1.32$). Then we compared the average difference between an individual's long-term partners with the expected average difference using a permutation test. We also evaluated differences between partners who had children with the participants (fathers) and other long-term partners (non-fathers) using permutation tests and mixed-effect models. Our results revealed that women choose long-term partners consistently with respect to all types of characteristics. Although effect sizes for the individual characteristics were rather weak, maximal cumulative effect size for all characteristics together was high, which suggests that relatively low effect sizes were caused by high variability with low correlations between characteristics, and not by inconsistent mate choice. Furthermore, we found that despite some differences between partners, fathers of participants' child(ren) do fit their 'type'. These results suggest that mate choice may be guided by relatively stable but to some degree flexible preferences, which makes mate choice cognitively less demanding and less time-consuming. Further longitudinal studies are needed to confirm this conclusion.

Keywords: repeatability, intraindividual variability, motherhood, stability of preferences, sexual selection, mating behavior, female preferences

INTRODUCTION

Human mate choices are influenced by various sociodemographic, physical, and psychological characteristics of a prospective partner. Majority of research on absolute partner preferences focuses on what is considered attractive across various individuals (see e.g., Buss, 1989; Regan et al., 2000). This line of research yielded evidence on high agreement with respect to attractiveness both within and across cultures ($r > 0.90$) (see a meta-analysis, Langlois et al., 2000). Agreement in attractiveness assessment between individual raters is, however, much lower ($r > 0.50$) (see a meta-analysis, Feingold, 1992). It seems therefore that despite a strong general consensus in attractiveness assessments in general, there exists a substantial variability between individual preferences (Hönekopp, 2006). This interindividual variability may be due to relative partner preferences (e.g., based on own characteristics and experiences, Figueredo et al., 2006; see a review, Štěrbová and Valentová, 2012). It is also possible, moreover, that an individual's partner preferences also change over time (Kościński, 2010).

In non-human animals, it has been found that individual consistency of female mate preferences is rather low and context-dependent, because it varies depending on females' age, environment, and conditions (Cotton et al., 2006; Jennions and Petrie, 2007; Bell et al., 2009). In humans, ontogeny of mate preferences has been studied mostly cross-sectionally (Brumbaugh and Wood, 2013; Boothroyd and Vukovic, 2018). This approach revealed differences among various age groups which were due to changes in hormonal levels, personal development, and the like (Kościński, 2011), but it did not track intraindividual variation in preferences in a longitudinal fashion. Kościński (2010) tested individual consistency of facial attractiveness assessment and found that self-correlation of women's assessment was approximately 0.78, which means that about 40% ($1 - 0.78^2$) of individual variation in attractiveness rating varies over time. To sum up, existing evidence suggests that preferences can change over time with age and reproductive stage of life, and that they can change in reaction to current circumstances (Rosenthal, 2017).

In short, it has been established that over time, mate preferences are to some degree plastic, but research of intraindividual stability in real mate choice in humans is sparse. To the best of our knowledge, only three studies so far tested individual consistency of mate choice (Eastwick et al., 2017; Newman et al., 2018; Štěrbová et al., 2018). They found consistency in preferences for eye color (Štěrbová et al., 2018; but cf. Newman et al., 2018), hair color (Štěrbová et al., 2018), attractiveness, masculinity, vitality, depression, delinquency, religiosity, educational aspirations, self-esteem, and intelligence (Eastwick et al., 2017). It is important to note, however, that the effect sizes were rather low. Nevertheless, it can be assumed that mate choice is affected by a vast array of demographic, physical, and personality traits.

In the present study, we tested individual consistency of mate choice with respect to traits that play an important role in human mating context and could therefore have a substantial impact on reproduction (Little, 2015). To wit, it is possible that different

characteristics are valued in non-reproductive as opposed to reproductive relationships, that is, that different characteristics result in direct versus indirect benefits to offspring (Boothroyd and Vukovic, 2018).

The following is a list of characteristics we followed with respect to stability of individual mate preference in women:

Tallness

Existing research suggests that women tend to choose partners who are tall (Hensley, 1994) and, in particular, taller than themselves (see meta-analysis, Pierce, 1996). This may be due to a link between body height and health (Christensen et al., 2007), and/or height and career success (Judge and Cable, 2004).

Body Form

Preferences for body shape and weight may reflect environmental variation in food availability (Anderson et al., 1992), but also serve as cues to an individual's social and economic status (Sarlio-Lahteenkorva and Lahelma, 1999). In general, optimal body mass index is perceived as attractive (Tovée et al., 1999). On top of that, metabolically expensive physical features, such as muscularity, are supposed to be attractive to females because they advertise that energy gathered from the environment could be converted to reproduction-related activities (Kaplan and Gangestad, 2004). Some studies found that women prefer muscular, but not too muscular, men (Dixson et al., 2003; Frederick and Haselton, 2007). In general, research supports the inverted-U hypothesis of masculine traits (Frederick and Haselton, 2007). These ambiguous results could be explained by personality characteristics associated with masculinity, such as higher dominance but also lower honesty, cooperativeness, emotionality, and parental qualities (Perrett et al., 1998; Boothroyd et al., 2008). Some studies thus found female preference for masculinity (Cunningham et al., 1990), whereas other research found preference for femininity in males (Perrett et al., 1998). Similarly, both hirsuteness and beardedness are sexual dimorphic traits. As in masculinity, evidence regarding female preferences is mixed (see for review, Dixson and Rantala, 2015), which could be due to association between beards and body hair on the one hand and perceived dominance and aggressiveness on the other hand (Puts, 2010).

Eye and Hair Color

Research shows that eye and hair color play an important role in some human populations (Frost, 2006; White and Rabago-Smith, 2011) because they can affect perceived trustworthiness (Kleisner et al., 2013), dominance (Kleisner et al., 2010), attractiveness (Laeng et al., 2006), and health status (Frost et al., 2017).

Personality

Last but not least, it has been established that cross-culturally, some personality traits likewise play an important role in mate choice (Buss and Barnes, 1986). It has been shown that both men and women desire partners who score high on Agreeableness, Openness (Botwin et al., 1997), and Emotional Stability (Conroy-Beam et al., 2015). These characteristics contribute to cooperation

and altruism (Jensen-Campbell and Graziano, 2001), and thereby have a positive impact on the couple's reproductive success (Buss, 1991).

The main aim of our study was to examine individual consistency of mate choice in women. Specifically, we tested whether women repeatedly choose long-term partners with particular demographic, physical, and personality characteristics. In short, we tested intraindividual variability of female mate choice. Consistency of mate choice was measured by several methods (by consistency index, percentage of variance in partners' trait values accounted for by the respondent, and by correlations). Effect sizes were estimated by stepwise randomization effect size assessment and stepwise estimation of shared effect size. Only mothers were recruited for the study because from an evolutionary perspective, the most important partner is the father of a woman's child or children. We have therefore tested mutual similarity among all of women's long-term (ex-)partners and tried to find whether the partner with whom they had a child or children is different from those partners with whom they did not reproduce.

MATERIALS AND METHODS

All procedures followed were in accordance with ethical standards of the relevant committee on human experimentation and with the Helsinki Declaration. The study was approved by the Institutional Review Board of Charles University, Faculty of Sciences (Approval number 2016/23). All participants were informed about the goals of the study and approved their participation by clicking button 'I agree' below the informed consent. Written informed consent was not obtained because the study was conducted online.

Participants

Respondents

Respondents were recruited via social sites, such as Instagram and Facebook, and websites aimed at mothering, e.g., babyweb.cz and emimino.cz, via flyers distributed to gynecology offices and dormitories, and by emails sent to respondents from our earlier studies. The initial sample consisted of 1,331 individuals. We analyzed only data from women who met the following criteria: (i) age between 18 and 45 years, (ii) heterosexual orientation (Kinsey scale < 3), (iii) had at least two long-term partners (defined as committed relationship that is believed to have future prospects), (iv) shared household with their biological father until at least 12 years old (this study is part of a larger research aimed at the imprinting-like effect).

The final sample consisted of 537 respondents (mean age = 29.14, $SD = 6.281$, median = 29, min = 19, max = 45). Information provided by each of these respondents was used in at least one analysis. All respondents together had a total of 1,599 partners (partners per respondent: mean = 2.98, $SD = 1.32$, median = 3, min = 2, max = 10). The mean length of relationship was 5.07 years ($SD = 4.99$, median = 3.17, min = 0, max = 27). When fathers and non-fathers were analyzed separately, mean length of relationship between these two categories of partners

differed (mean length of relationship with fathers in years = 8.44, $SD = 5.65$, median = 7.416, min = 0.25, max = 27, mean length of relationship with non-fathers in years = 2.61, $SD = 2.32$, median = 2, min = 0, max = 20).

In many cases, respondents did not report all 21 characteristics about all partners, which prevented us from including all respondents and all partners in all tests of mate choice consistency. Respective sample sizes did not differ substantially (the number of respondents: mean = 482.8, $SD = 21.9$, median = 481, min = 435, max = 516; Number of partners with known information: mean = 1.388, $SD = 64.3$, median = 1.376, min = 1.236, max = 1.491) and are all reported in the **Appendix**.

Measurements

Respondents reported a total of 21 characteristics (3 demographic and 13 physical characteristics, and 5 personality traits). Since some of these characteristics can change within a short period of time (e.g., beardedness), respondents were asked to indicate characteristics as they remember them from the time when the relationship was established.

Of demographic characteristics, they were asked to report the size of their partners' and fathers' residence (1 – metropolis, 2 – city, 3, town, 4 – village), education level (1 – elementary school, 2 – high school, 3 – college, 4 – university), and age difference between themselves and their long-term partners (in months; negative numbers indicate that a woman is older than her partner).

Physical characteristics were reported by selecting the relevant image stimuli of eye color (gray, blue, green, brown, and black), hair color (9 shades varying from light blond to black), facial masculinity (five images varying from low to high masculinity) (Johnston, 2006), beardedness (four images varying from clean shaven to fully bearded) (Dixson and Brooks, 2013), muscularity (six images varying from low to high muscularity) (Frederick and Haselton, 2007), relative height (six images varying from man-taller pattern to women-taller pattern) (Pawlowski, 2003), body mass index (six images varying from low to high BMI) (Allen et al., 2003), hirsuteness (five images varying from a low to a high level of hirsuteness) (Dixson et al., 2010), leg to body ratio (LBR) (five images varying from relatively short to long legs) (Swami et al., 2006). Further, respondents indicated their partners' body weight (in kilograms), body height (in centimeters), and overall masculinity and attractiveness (using a 7-point verbally anchored Likert scale, ranging from 'under average' to 'above average').

To assess personality characteristics, we used the Ten-Item Personality Inventory (TIPI) (Gosling et al., 2003), which maps five domains: Emotional Stability, Extraversion, Openness, Agreeableness, and Conscientiousness. Responses for each item were indicated on a 7-point Likert scale ranging from 'strongly disagree' to 'strongly agree.' We used a method of translation and back-translation into the Czech language.

PROCEDURE

The test was administered online using the Qualtrics platform. At the outset, respondents were asked to confirm their

informed consent. In order to examine consistency of mate choice, respondents described their all (ex-)partners using image stimuli to assess their physical traits, indicated their demographic characteristics, and answered questions in a personality assessment questionnaire. They indicated only those characteristics they clearly remembered, otherwise they were asked to skip the question. A total of 206 (38.36%) out of 537 individuals involved in the analysis failed to indicate at least one of their partners' characteristics. To make the process easier on the participants, they indicated the names or assigned nicknames to all (ex-)partners at the beginning of questionnaire. These (nick)names were then displayed with each question, meaning the respondent did not have to remember the order of (ex-)partners and we could be reasonably certain that for each partner, we know all the characteristics the participant remembers. This study was part of a wider research, which is why respondents also answered questions pertaining to their fathers. The whole procedure took approximately 80 min.

ANALYSES

All analyses described below were conducted using the R 3.5.1 software. The code is available at: <https://github.com/costlysignalling/Mate-choice-consistency-2>.

Consistency Evaluation

Average difference between respondents' partners ($\bar{\Delta}$) served as a measure of mate choice consistency. Larger differences between respondents' partners indicate a more diverse set of partners and a lower mate choice consistency.

To assess the consistency of mate choice, we used a procedure similar to consistency index described in an earlier study (Štěrbová et al., 2018). Since the original consistency index views qualitative character states only in terms of identity (1) or difference (0) between pairs of respondent's partners with respect to a particular character state, we used the average difference between respondents' partners ($\bar{\Delta}$) as a parametric equivalent of consistency index.

First, we assessed the average difference between a pair of partners separately for each respondent. For example, if only two long-term partners were reported, one had extroversion value 11 and the other 13, the average difference between them was 2 (i.e., $13 - 11$). When four long-term partners were reported, their extroversion values could be as diverse as 5, 10, 7, and 14. In such a case, we calculated mutual differences for every possible pair of partners ($10 - 5 = 5$, $7 - 5 = 2$, $14 - 5 = 9$, $10 - 7 = 3$, $14 - 10 = 4$, $14 - 7 = 7$) and then computed the average, i.e., $(5 + 2 + 9 + 3 + 4 + 7)/6 = 5$. This average value characterizes a woman's mate choice consistency with respect to a particular trait. These individual values were later averaged across all respondents in the sample to evaluate overall mate choice consistency [in this short example, that value would be calculated as $(2 + 5)/2 = 3.5$]. This way, we ensured that every woman contributed to population consistency equally, i.e., regardless of the number of her long-term partners.

Populational average difference between all partners of an individual could thus be expressed as:

$$\bar{\Delta} = \sum_{i=1}^n \sum_{j=1}^{p_i-1} \sum_{k=j+1}^{p_i} \frac{|t_{ij} - t_{ik}|}{p_i(p_i - 1)/2n}$$

where $\bar{\Delta}$ indicates the populational average difference between partners of an individual, p_i the number of partners of i -th individual, t_{ix} trait value of x -th partner of i -th individual, and n the number of respondents.

Permutation Test of Mate Choice Consistency

Subsequently, we compared the observed average difference between an individual's partners ($\bar{\Delta}$) with the distribution of expected $\bar{\Delta}$ in a population with random pairing. A permutation test was executed to obtain the equivalent of a unidirectional test p -value.

We assigned partners to respondents randomly while maintaining the number of partners each respondent actually reported. This was done for each trait separately. We generated 10,000 such random populations and calculated the $\bar{\Delta}$ for each one. This yielded the distribution of $\bar{\Delta}$ for a random pairing.

We assessed the proportion of $\bar{\Delta}$ in random permutations which were smaller than the observed value of $\bar{\Delta}$. This gave us the equivalent of one-tailed test p -value, which indicated whether people were indeed significantly more consistent in their mate choice than one would expect if the choice were random.

Stepwise Randomization Effect Size Assessment

This procedure allowed us to estimate the proportion of partners that have to be switched between respondents in order to lower the mate choice consistency to the expected level. This measure can range between 0% (observed consistency is lower than or equal to the expected consistency and no partners need to be switched) and 50% (see the example below).

We calculated the effect size attributable to consistency of mate choice using a stepwise randomization test. In this test, the observed $\bar{\Delta}$ is gradually elevated by random relocation of one partner at a time until the expected value of $\bar{\Delta}$ is reached (the procedure is described in detail in Štěrbová et al., 2018). The resulting percentage indicates the proportion of partners that needs to be switched among participants until one arrives at a $\bar{\Delta}$ that would be expected in a random pairing. This was done 10,000 times for each trait. The mean value is reported as the estimated effect size together with a 95% confidence interval of this measure calculated as a 2.5–97.5% quantile of these 10,000 permutation-yielded percentages. It should be noted that though expressed as a percentage, this effect size is not identical to the proportion of explained variance. Maximal effect size in terms of percentage of partners that need to be relocated is about 50% because after relocating one half of all partners, one necessarily starts approaching the initial configuration again.

For instance, imagine a hypothetical dataset where every respondent reports 2 partners who have the same, either brown

or blue, eye color. The observed $\bar{\Delta}$ in such a population is 0. The maximal $\bar{\Delta}$ is observed when data are permuted so that each individual has one brown-eyed and one blue-eyed partner. This state is achieved when 50% of partners switch places. Relocation of one more blue-eyed or black-eyed partner would necessarily result in coupling with another partner of the same eye color, which would lower the overall $\bar{\Delta}$.

Other Measures of Mate Choice Consistency

Several non-permutation methods based on correlational approach had been proposed to tackle the mate choice consistency problem. These methods are to some extent equivalent to the proportion of partners to be switched between respondents (see section “Stepwise Randomization Effect Size Assessment”), but they are expressed either as a correlation coefficient or as the proportion of explained variance (values between 0 and 1 or 0 and 100%, respectively).

As suggested earlier (Eastwick et al., 2017), the percentage of variance in partners’ trait values accounted for by the respondent (i.e., the metric conceptually identical to the Intraclass correlation coefficient) can be used as a measure of mate choice consistency in parametric variables. We calculated this measure as well to enable a comparison with stepwise randomization effect size. We treated respondent identity (ID) as a random factor in a mixed-effect model (using lmer function within the lmerTest package). The statistic of interest was the random variance estimate for respondent ID divided by total variance. Reasonable benchmarks of this variance estimate were outlined at 10% for meaningful, 20% for a medium-sized effect, and 30% for a large effect (Kenny et al., 2006).

We have also calculated a simple Pearson correlation coefficient between two vectors of partners’ trait values. Every possible pair of partners of the same individual was treated as a unit of analysis. Individuals who had more partners therefore contributed to the overall coefficient disproportionately. Comparisons between this measure and more rigorously estimated effect sizes described above might indicate, however, that this is not necessarily a problem.

Pearson correlation coefficients between these three effect size measurements were calculated to demonstrate the equivalence of these measures. Additionally, we evaluated a linear model of dependence between the explained variance and the proportion of partners that needs to be switched between partners. This provided a useful tool for future comparisons with results on mate choice consistency that would use different approaches to effect size reporting.

Stepwise Estimation of Shared Effect Size

Elaborating on the permutational effect size estimation (see section “Stepwise Randomization Effect Size Assessment”), we can assess shared effect size between mate choice consistency along two non-independent variables. Permutational effect size is expressed in the proportion of partners that needs to be switched between respondents. Shared effect size is the proportion of

partners switched in two seemingly independent estimates of consistency effect size where correlation between variables is taken into account. If one switches 16% of partners to reach the expected consistency in body weight, and then another 6% are switched to avoid also consistency in body height, one could claim that 22% of all partners need to be switched to avoid non-random consistency in both height and weight. Going in the opposite direction, we relocate 11% to avoid consistency in height and then another 11% to avoid consistency in weight. Since the sum of residual effect sizes ($6\% + 11\% = 17\%$) is lower than the sum of simple effect sizes ($16\% + 11\% = 27\%$), one can assume non-independence between these variables and calculate a shared effect size. This ‘overlap’ is missing in the sum of residual effect sizes 22–17% and present twice in the sum of simple effect sizes 27–22%, but in both cases, the resulting proportion is 5%. These shared effect sizes can be used to calculate the maximal cumulative effect size, i.e., the number of partners that need to be switched between individuals to avoid mate choice consistency on all characteristics.

The link between every pair of partners’ qualities was assessed in two ways, namely Pearson correlation coefficient with a single partner as a unit of analysis and shared effect size, which is equivalent to the abovementioned stepwise randomization effect size for a pair of variables (A and B). Here, the $\bar{\Delta}$ in variable A of empirical population is elevated by a stepwise reassignment of partners until the mean expected value of consistency with respect to A is reached. This rearranged population is then taken as a starting point and stepwise randomization effect size assessment is executed for variable B. The residual proportion of partners that need to be switched to avoid consistency in B is estimated after the effect of consistent mate choice with respect to A is eliminated. This is done 1,000 times to get the average residual effect size of B, and 1,000 times in the opposite direction to get the equivalent measure for A. Shared effect size is then calculated easily as $A + B = (A \cap \neg B) + (B \cap \neg A) + 2 \times (A \cap B)$, where \cap represents intersection, \cup unity, and \neg a set complement. This value was calculated for every pair of assessed variables.

Maximal cumulative effect size was then derived from pairwise shared effect sizes. Higher-order intersections were not estimated with permutation approach. It would have been possible, but extremely demanding with respect to computation time. Instead, we assumed that these intersections are proportional to the ratio of pairwise intersections. For example, if variables A, B, and C have effect sizes of 20, 15, and 10% partners to switch, and their shared effects are 10% for $A \cap B$, 4% for $A \cap C$, and 3% for $B \cap C$, it is assumed that segments $(A \cap C) \cap \neg B$, $(B \cap C) \cap \neg A$, and $A \cap C \cap B$ are in the same proportion as $A \cap \neg B$, $B \cap \neg A$, and $A \cap B$ (i.e., 10:5:10), and given that in the sum of $A \cap C$ and $A \cap B$ ($4 + 3 = 7$), segment $A \cap B \cap C$ appears twice, it follows that the final proportions will be 2% for $(A \cap C) \cap \neg B$, 1% for $(B \cap C) \cap \neg A$, and 2% for $A \cap B \cap C$. The unique contribution of C [$C \cap \neg (A \cup B)$] must equal 5% of partners and the total cumulative effect size ($A \cup B \cup C$) must be 30%. To minimize possible errors stemming from inaccuracy of the assumption of intersection proportionality, variables were added to the total cumulative effect size one at a time according to a criterion of

maximal unique contribution to the total effect size. First, we included variable A, which had the largest unique effect size, then we calculated for all other variables their unique contribution to the total effect size, selected the one which contributed the most, labeled it B, and included it in our calculation. For the next variable, we calculated its contribution to the union of A and B, added to the model the one with the largest unique contribution, and so on.

Growth of the unique contribution relative to the previous step of variable inclusion was a sign of accumulated error caused by inaccuracy of the assumption of intersection proportionality (i.e., in this step, the variable was rearranged back to high mate choice consistency). In each step, therefore, higher unique contributions were replaced by minimum values from contributions calculated in previous steps. This number represents the minimal possible contribution without allowing for a negative relationship between consistent mate choice along different variables. As a result, consistent mate choice in one variable or a union of variables could lead to inconsistent mate choice in another variable or variables. Maximal cumulative effect size was calculated as the total sum after the stepwise addition of all variables. The fact that a negative relationship between consistency on different variables was neglected is not problematic because the individual contributions still add up to the same total. Sacrifice of a consistent mate choice on one variable is compensated by an equivalent increase in the consistency along other variables. Therefore, although the order of unique contributions to overall consistency and their magnitude may be burdened by an error, the estimate of maximal cumulative effect size is sound and reliable.

Since a high number of assessed mate choice consistencies and their shared effects naturally leads to a substantial maximal cumulative effect size, the empirical level is contextualized with the expected maximal cumulative effect size in a population with random pairing. In this resampled population, the identity of partners and links between their qualities remained identical to the empirical data, but partners were scrambled among respondents so the total number of partners any respondent had remained unchanged.

Differences Between Fathers and Non-fathers

Differences between partners with whom the respondents had children ('fathers') and other former long-term partners ('non-fathers') were investigated along all 21 romantic partner qualities our study had followed. Mean values and variances of fathers and non-fathers were compared to reveal possible differences between these groups. Changes in $\bar{\Delta}$ after the exclusion of fathers were compared to expected changes in $\bar{\Delta}$ after the exclusion of random individuals to assess whether fathers were especially typical of given partner sets and elevated overall mate choice consistency, or exceptional within these sets, thus lowering overall mate choice consistency.

The $\bar{\Delta}$ was calculated first for the full partner set and then for a restricted sample where fathers were excluded. The observed difference between these two samples was compared with the

distribution of expected differences yielded by a permutation test. In each referential permutation, fathers were selected randomly from sets of partners provided by respondents. For instance, if a participant reported four long-term partners and two of them fathered at least one of her children, two individuals from this set were labeled as fathers and excluded in each permutation run (as expected, however, most respondents had children with only one partner). Two tailed p -value was calculated as a measure of significance of the difference between measured and expected changes in mate choice consistency after the exclusion of fathers. 10,000 permutation runs were executed for each variable. Where non-father $\bar{\Delta}$ was significantly higher than expected, fathers were highly typical (or intermediate) representatives of woman's partners. Where it was lower than expected, fathers were rather exceptional individuals within the sets of partners and measured consistency of mate choice was higher without them.

Yet even if fathers fitted in partner sets without being either exceptionally typical or highly atypical, it was still possible that there are some differences between fathers and non-fathers along the assessed qualities. Mixed effect models were employed to calculate the probability of equality of group means. Mixed effect equivalents of Levene's test, where distance from the group mean is used as a response variable, were then used to investigate equality of variances between father and non-father groups, since it could be the case that even if the two groups do not differ in their means, the extreme or intermediate individuals just may not be the right 'father material.' We treated respondent ID as a random factor in all mixed effect models and used the lmer function lmer from the lmerTest package.

All independent sets of p -values reported in the result section were adjusted for multiple comparisons using the Benjamini-Hochberg procedure. Vectors of p -values calculated for the sets of 21 qualities were adjusted separately, while the p -values of correlations between qualities were adjusted together.

RESULTS

Mate choice consistency was higher than expected in all assessed qualities except for facial masculinity and beardedness. Difference between observed and expected consistency was statistically significant in most qualities, but effect sizes differed substantially. While consistency of mate choice in residence or weight was substantial, it was only medium-sized or small with respect to hair or eye color. Complete results are summarized in **Table 1** and **Figure 1**.

The average effect size was highest in demographic variables, but none of the pairwise comparisons between groups of variables (demographic, physical, and psychological) was statistically significant ($p > 0.1$). Permutation test results are visualized in **Figure 1**. All sample sizes and descriptive statistics of all variables are listed in the **Appendix**. The different estimates of effect size were highly correlated. The proportion of males who had to be relocated between respondents correlated with the variance accounted for by the respondent at 0.93, whereby a linear model of relationship between these two measures supports the idea that the latter is approximately double of the former. The slope

TABLE 1 | Mate choice consistency: complete results.

	Observed $\bar{\Delta}$	Expected $\bar{\Delta}$	SD (of expected $\bar{\Delta}$)	<i>p</i> -value	Proportion of partners to switch (%) (95% CI)	Respondent variance (%)	Pearson <i>r</i>
Residence	0.72	1.23	0.03	<0.001	19.48(18.03,20.9)	43.58	0.42
Education	0.72	0.87	0.02	<0.001	7.17(6.45,7.88)	19.24	0.18
Weight	8.82	11.31	0.24	<0.001	16.21(14.04,18.39)	33.77	0.31
Height	6.3	7.42	0.16	<0.001	11.28(9.53,13.06)	24.56	0.22
Age difference	44.05	49.44	1.13	<0.001	10.49(8.41,12.7)	13.97	0.14
Attractiveness	1.23	1.44	0.03	<0.001	8.28(7.26,9.36)	16.49	0.19
Masculinity	1.26	1.4	0.03	<0.001	5.51(4.61,6.44)	9.71	0.11
Eye color	1.28	1.35	0.03	0.026	2.76(2.17,3.43)	4.31	0.05
Hair color	2.54	2.74	0.07	0.001	4.81(3.89,5.77)	10.17	0.12
Facial masculinity	1.66	1.46	0.04	1.000	0.07(0.07,0.07)	0	-0.05
Beardedness	0.67	0.67	0.02	0.575	0.07(0.07,0.07)	6.83	0.08
Muscularity	1.01	1.03	0.03	0.181	1.27(0.87,1.67)	3.43	0.05
BMI	1.25	1.35	0.03	0.002	4.23(3.5,4.99)	11.87	0.17
Relative height	1.18	1.5	0.04	<0.001	11.99(10.75,13.3)	30.58	0.34
Hirsuteness	1.41	1.48	0.04	0.055	2.56(1.93,3.26)	10.31	0.12
Leg to body ratio	1.21	1.35	0.03	<0.001	5.93(4.98,6.89)	13.52	0.12
Extraversion	3.12	3.36	0.08	0.002	5.25(4.2,6.41)	9.32	0.12
Agreeableness	2.94	3.14	0.07	0.003	5.07(4.1,6.11)	3.55	0.06
Conscientiousness	3.68	3.92	0.09	0.006	4.68(3.62,5.76)	5.08	0.09
Emotional stability	3.29	3.57	0.08	0.001	5.84(4.69,7.08)	8.1	0.12
Openness	2.91	3.21	0.07	<0.001	6.94(5.77,8.16)	9.27	0.11

Results of permutation tests of mate choice consistency

comparing observed average difference between partners of an individual ($\bar{\Delta}$) with the distribution of expected $\bar{\Delta}$ values in population with random pairing

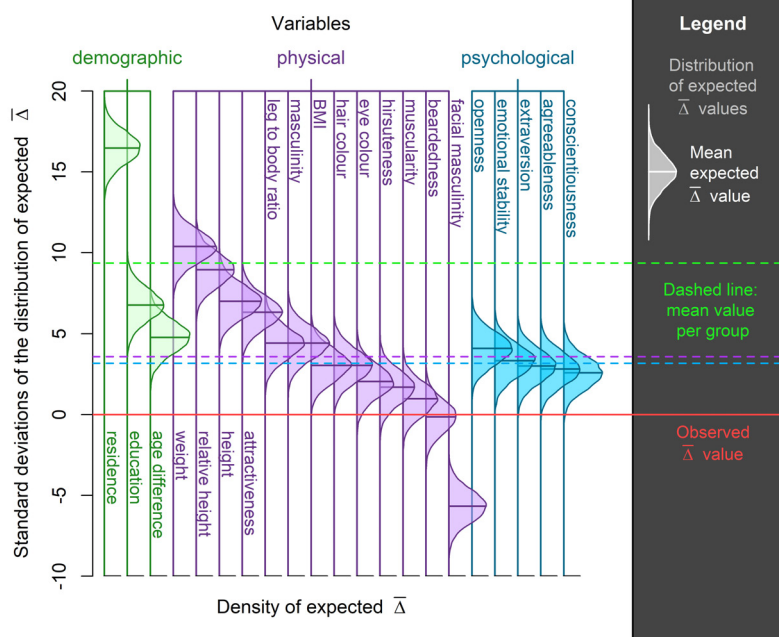
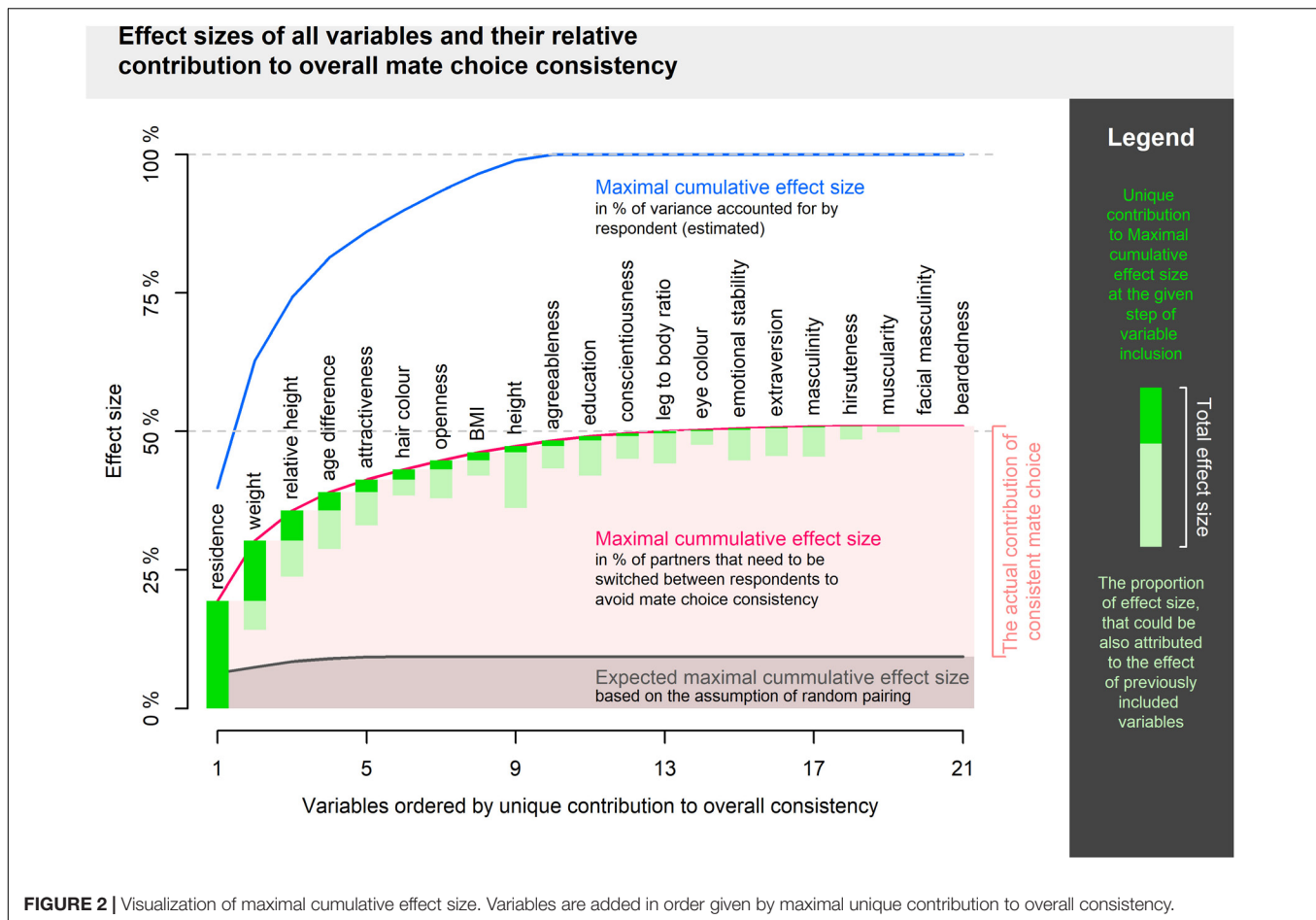


FIGURE 1 | Visualization of permutation tests of mate choice consistency centered around observed $\bar{\Delta}$ and normalized along the SD of expected $\bar{\Delta}$ distribution. Difference between the observed and expected value is expressed in standard deviations from the expected value distribution. The higher the bell curve above the Observed $\bar{\Delta}$ value, the higher the actual mate choice consistency. Bell curve below Observed $\bar{\Delta}$ value indicates a trait where the observed mate choice was less consistent than expected.

TABLE 2 | Relations between investigated qualities of romantic partners expressed in shared effect sizes and Pearson correlations.

	Residence	Education	Weight	Height	Age difference	Attractiveness	Masculinity	Eye Color	Hair Color	Facial masculinity	Beardedness	Muscularity	BMI	Relative height	Hirsuteness	Leg to body ratio	Extraversion	Agreeableness	Conscientiousness	Emotional stability	Openness
Residence	19.4%	-0.26*	-0.02	-0.06*	-0.01	-0.03	-0.04	-0.07*	-0.08*	-0.04	-0.04	-0.03	-0.01	-0.02	0.02	-0.07*	-0.01	0.03	0.00	0.01	-0.11*
Education	3.2%	7.1%	0.06	0.12*	-0.07*	0.04	0.05	0.00	0.04	0.01	0.02	0.03	0.04	-0.03	0.09*	0.10*	-0.07*	0.13*	0.27*	0.16*	0.10*
Weight	5.3%	1.8%	16.1%	0.48*	0.17*	-0.02	0.26*	-0.01	0.05	0.09*	0.16*	0.24*	0.59*	-0.25*	0.27*	0.02	0.04	-0.02	0.08*	0.05	0.03
Height	4.0%	1.6%	5.0%	11.2%	0.01	0.08*	0.14*	-0.03	0.02	0.19*	0.05	0.07*	0.04	-0.58*	0.03	0.44*	0.03	0.03	0.06	0.09*	0.04
Age difference	3.5%	1.5%	3.3%	2.5%	10.3%	-0.03	0.13*	-0.01	0.08*	0.15*	0.14*	0.07*	0.18*	0.06	0.14*	-0.04	0.05	-0.03	0.03	-0.01	-0.01
Attractiveness	2.7%	1.4%	2.5%	2.1%	1.7%	8.3%	0.45*	0.02	0.09*	0.19*	0.09*	0.28*	-0.13*	-0.05	0.00	0.07*	0.19*	0.17*	0.10*	0.18*	0.19*
Masculinity	1.8%	0.8%	2.2%	1.3%	1.4%	2.1%	5.4%	0.01	0.05	0.28*	0.16*	0.50*	0.18*	-0.10*	0.18*	-0.04	0.16*	0.11*	0.13*	0.19*	0.18*
Eye color	1.1%	0.3%	0.7%	0.6%	0.6%	0.4%	0.2%	2.7%	0.39*	0.03	0.04	-0.03	-0.02	-0.01	0.04	-0.03	0.03	0.00	-0.07*	-0.03	0.00
Hair color	1.2%	0.7%	1.0%	0.9%	0.3%	0.7%	0.2%	0.6%	4.7%	0.08*	0.15*	0.03	0.02	0.01	0.17*	0.01	0.02	-0.01	-0.03	-0.02	0.02
Facial masculinity	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.19*	0.31*	-0.08*	-0.12*	0.16*	0.15*	0.04	-0.02	0.00	0.06	0.06
Beardedness	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.13*	0.14*	-0.03	0.34*	-0.01	0.07*	0.06	0.02	0.08*	0.12*
Muscularity	0.4%	0.1%	0.6%	0.2%	0.1%	0.5%	0.5%	0.1%	0.3%	0.0%	0.0%	1.2%	0.19*	0.01	0.17*	-0.08*	0.09*	0.10*	0.12*	0.15*	0.13*
BMI	0.6%	0.1%	1.7%	0.3%	0.2%	0.5%	0.6%	0.0%	0.1%	0.0%	0.0%	0.4%	4.2%	0.01	0.30*	-0.25*	0.02	-0.01	0.08*	0.00	-0.01
Relative height	4.1%	1.4%	3.7%	4.9%	2.8%	2.3%	1.2%	0.6%	0.8%	0.0%	0.0%	0.3%	0.4%	11.9%	0.01	-0.39*	0.00	-0.02	-0.05	-0.06	0.00
Hirsuteness	0.9%	0.2%	1.0%	0.3%	0.5%	0.4%	0.4%	0.0%	0.1%	0.0%	0.0%	0.2%	0.4%	0.4%	2.4%	-0.10*	0.01	0.03	0.08*	0.01	0.03
Leg to body ratio	2.0%	0.7%	1.6%	2.6%	1.3%	1.1%	0.8%	0.1%	0.3%	0.0%	0.0%	0.2%	0.4%	2.4%	0.5%	5.8%	-0.05	0.02	0.03	0.04	0.02
Extraversion	1.9%	0.7%	1.7%	1.3%	0.9%	1.2%	0.5%	0.3%	0.5%	0.0%	0.0%	0.1%	0.2%	1.2%	0.4%	0.6%	5.1%	0.09*	-0.09*	0.17*	0.38*
Agreeableness	1.8%	0.5%	1.6%	0.7%	0.2%	0.9%	0.9%	0.1%	0.2%	0.0%	0.0%	0.2%	0.2%	0.5%	0.4%	0.5%	1.0%	5.0%	0.33*	0.48*	0.22*
Conscientiousness	1.7%	0.7%	1.0%	0.8%	0.7%	0.9%	0.9%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.6%	0.3%	0.3%	0.9%	1.2%	4.6%	0.38*	0.08*
Emotional stability	2.2%	0.9%	2.0%	1.4%	1.0%	1.1%	0.8%	0.2%	0.3%	0.0%	0.0%	0.3%	0.3%	0.9%	0.5%	0.5%	1.2%	2.0%	1.3%	5.8%	0.19*
Openness	2.5%	1.0%	1.6%	1.2%	1.2%	1.3%	0.8%	0.2%	0.5%	0.0%	0.0%	0.3%	0.3%	1.4%	0.3%	0.4%	1.6%	1.2%	1.0%	1.5%	6.9%

Indicated on the diagonal is the total effect size in terms of proportion of partners who have to be switched between respondents to avoid non-random consistency of mate choice along a given variable (see section "Stepwise Randomization Effect Size Assessment"). Shared effect sizes, i.e., proportions of partners that one switches in both seemingly independent estimations of consistency effect size due to correlation between variables (see section "Stepwise Estimation of Shared Effect Size"), are located below the diagonal. Correlations between variables with partner as a unit of analysis are displayed above the diagonal. *P*-values are adjusted for multiple comparisons with Benjamini-Hochberg correction procedure. Significance levels after this correction are indicated: * $p < 0.05$.



in the model where respondent-attributable variance regressed on the proportion of partners to relocate was 2.08 (95% CI = 1.72–2.45) with minimal (not significantly different from 0) intercept of -0.18 (95% CI = -3.19–2.83). Results yielded by the simple Pearson correlation correlated at 0.91 with the percentage of partners to relocate and at 0.98 with respondent-attributable variance. All of these measures can be thus treated as functionally equivalent.

Links between pairs of partners' qualities are summarized in **Table 2**. In total, 103 out of 210 correlations were significant even after Benjamini–Hochberg correction for multiple comparisons. Maximal cumulative effect size was 50.95% (expressed in the proportion of partners to switch between individuals). The first 10 variables ordered according to their unique contribution starting with the highest (residence, weight, relative height, age difference, attractiveness, hair color, openness, BMI, height, agreeableness, in this order) explained 48.30% of partner assignment. The other 11 variables contributed little (their unique contributions were less than 1%) or not at all (after the inclusion of all other variables, facial masculinity and beardedness failed to show any positive numbers). Full results are visualized in **Figure 2**.

Reaching maximal possible effect size suggests that adding yet other variables to a similar model of cumulative consistency

would add little to our current sum. On the other hand, it is conceivable that one might select precisely those variables which are not intercorrelated and explain a majority of mate choice consistency in just a handful independent dimensions. In theory, complex interaction patterns may lead to an even higher cumulative effect size since 50% of partners to relocate as an effect size limit applies to a single variable with two levels and represents the difference between maximal and minimal consistency (i.e., not maximal and expected). The high proportion of significantly correlated pairs of variables (49%), does, however, fit well within the impression of a substantial redundancy in our model.

Permutation test of changes in mate choice consistency revealed that fathers are significantly exceptional amongst participants' long-term partners in beardedness, muscularity, hirsuteness, extraversion, and openness. The average $\bar{\Delta}$ without these individuals was lower than the $\bar{\Delta}$ in permutation runs where an equivalent proportion of random partners (i.e., fathers and non-fathers) was excluded. Fathers were not significantly typical long-term partners in any of the assessed qualities. Complete results of these tests are summarized in **Table 3** and visualization is provided in **Figure 3**.

In qualities where fathers were indicated as exceptional individuals (except for extraversion), mean trait values differed between fathers and non-fathers, while variances differed in

TABLE 3 | Permutation test of father exceptionality, complete results.

	Change in $\bar{\Delta}$ when fathers are excluded	Expected change in $\bar{\Delta}$	SD (of expected change in $\bar{\Delta}$)	p-value
Residence	−0.01	0.02	0.03	0.717
Education	0.01	−0.01	0.03	0.725
Weight	0.08	0.16	0.28	0.925
Height	−0.18	−0.10	0.19	0.865
Age difference	−2.96	−0.32	1.42	0.176
Attractiveness	0.03	0.02	0.04	0.984
Masculinity	−0.07	0.01	0.04	0.141
Eye color	−0.01	−0.01	0.04	0.978
Hair color	−0.02	−0.08	0.08	0.755
Facial masculinity	−0.05	−0.05	0.04	0.978
Beardedness	−0.15	−0.04	0.03	0.002
Muscularity	−0.14	−0.03	0.03	0.002
BMI	0.04	0.03	0.04	0.925
Relative height	−0.01	0.01	0.04	0.834
Hirsuteness	−0.24	−0.11	0.05	0.024
Leg to body ratio	−0.08	−0.03	0.04	0.409
Extraversion	−0.29	−0.01	0.10	0.024
Agreeableness	−0.13	0.04	0.09	0.176
Conscientiousness	−0.24	−0.05	0.12	0.191
Emotional stability	−0.29	−0.09	0.10	0.158
Openness	−0.42	−0.08	0.09	0.004

beardedness, muscularity, and hirsuteness. Fathers were more bearded, hairier, more muscular, and showed a higher openness to experience. These differences might explain the overall exceptionality of fathers except for extraversion. It seems that fathers are outliers within partner sets even where the group means and variances of father and non-father sets do not differ. Moreover, fathers lived in larger cities, had higher education, were heavier and taller (although relatively, their height was closer to the height of respondents), more attractive and masculine, had lighter eyes, darker hair, more masculine faces, and were more agreeable, conscientious, and emotionally stable than non-fathers.

Group variances differed in several qualities. Fathers were significantly more variable than non-fathers with respect to age difference from the respondent and less variable in attractiveness, masculinity (general and facial), BMI, conscientiousness, and agreeableness. It seems that along these variables, either or both of the extremes are not the right for the ‘father material’. A graphic overview which compares densities that indicate differences between group means and variances is presented in **Figure 4**. Complete results in a textual form are listed in **Table 4**.

DISCUSSION

The aim of this study was to examine consistency of mate choice with respect to a variety of demographic, physical, and personality characteristics. We found that women choose long-term partners consistently across all types of characteristics (demographic, physical, and personality), but consistency was not observed in all tested traits. We also investigated potential

differences in tested characteristics between long-term ex-partners and partner(s) with whom women had child(ren). Results revealed that fathers in general fit the women’s ‘type’, although differences between them and other (ex-)partners are not large. Our findings are in line with earlier research (Eastwick et al., 2017; Štěrbová et al., 2018 but cf. Newman et al., 2018), which found that people consistently choose partners with certain traits, although reported effect sizes were rather small.

Is there any potential advantage to having a ‘type’? We could assume that preference for a particular ‘type’ may facilitate mate choice decisions. In theory, the pool of potential partners is immense and in the most extreme case covers almost one half of adult human population on Earth. This theoretical pool is, of course, unrealistic, but even so, people do have many potential partners to actually choose from. In order to navigate this vast amount of opportunities, it may be useful to follow a certain direction in this multidimensional trait space of human characteristics. Preference for a certain ‘type’ would constrain the spectrum of potential choices and reduce the dimensionality of trait space. A systematic, ‘type-directed’ exploration of this multidimensional trait space would facilitate better orientation on the ‘mating market.’ In short, having a ‘type’ means that women need not create new preferences always anew and based on random choices, i.e., it precludes them from jumping unsystematically across the vast dimensionality of trait space. A ‘type’ should not be viewed as a rigid attractor but rather as a polarizing filter which canalizes the selection of optimal partner. A ‘type’ is thus not a target in itself but rather the means by which a goal can be reached (and, e.g., an appropriate partner for reproduction, a father, found). This is why fathers do not fully correspond to a typical

Results of permutation tests of father exceptionality

tracing changes in mate choice consistency measured by $\bar{\Delta}$ after exclusion of fathers with the distribution of expected changes after exclusion of random partners

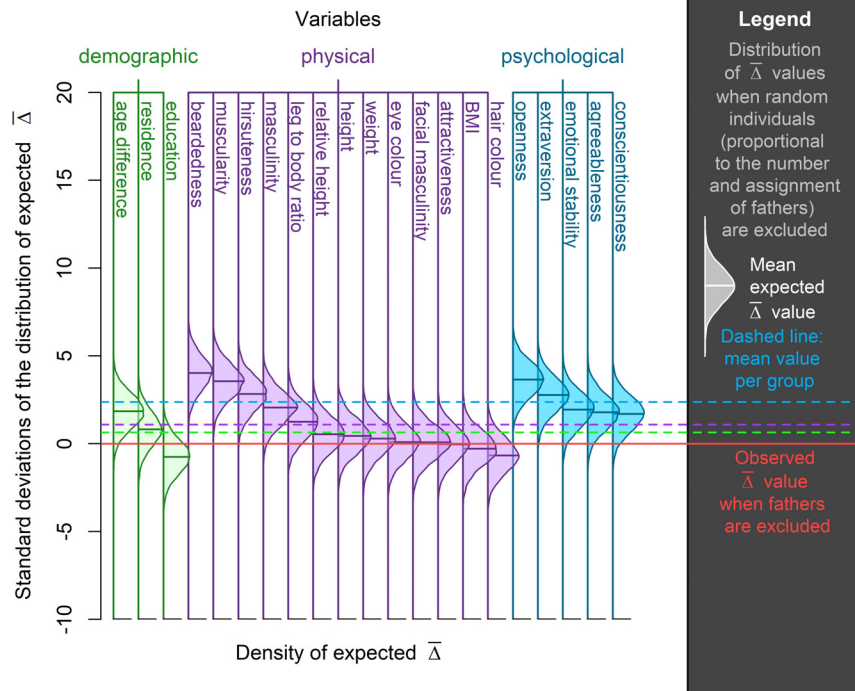


FIGURE 3 | Visualization of permutation tests of father exceptionality centered around the observed $\bar{\Delta}$ when fathers were excluded from the sample of partners and normalized along the SD of expected $\bar{\Delta}$ distribution in such a situation. Difference between observed and expected values is expressed in standard deviations of expected value distribution. The higher the bell curve above the observed $\bar{\Delta}$ value, the more exceptional were the fathers among the long-term partners of an individual. Bell curve below the observed $\bar{\Delta}$ value indicates a trait where fathers were more typical representatives of an individual's long-term partners.

partner and show some, however, small, deviation from the type.

This setup of optimal partner preferences may be beneficial. Mate choice not guided by such relatively stable but to some degree flexible preferences would be much more cognitively demanding and time-consuming. What remains unclear, however, is when and how are these preferences established. One of such mechanisms could be the imprinting-like effect (parent-partner similarity) or homogamy (self-similarity) (see Štěrbová et al., 2018). Moreover, parent-partner similarity can be promoted by emotional closeness with a parent during childhood (Saxton, 2016). Some plasticity of preferences may be adaptive also because it helps individuals adjust their preferences according to the current situation (e.g., ecological circumstances, inner state, their own characteristics which vary over time, experiences). From an evolutionary perspective, variation in mate preferences is important for speciation and diversification (Rodríguez et al., 2013). Species can adapt to changing circumstances by adjusting their mate choice. One might assume that learning would decrease the consistency of mate choice, that one would, for instance, choose a partner with characteristics different from an earlier partner because of negative experiences. On the other hand, mate choice is a mostly

non-conscious process, which implies that partner preferences are not easily modulated by experience. Our findings support these assumptions, because we found that women have a 'type' and choose partners who fit it.

One can only speculate whether mate choice when reproduction is in question differs from earlier preferences, i.e., preferences in a non-reproductive context. From an evolutionary perspective, the most important partner is the one with whom a woman will reproduce. This is why we tested whether fathers fit the women's 'type,' or rather whether fathers' characteristics differ from characteristics of the non-fathers.

Our results show that although consistency is found across all of woman's long-term partners, there are some notable differences between non-fathers and fathers. In particular, fathers disrupted consistency in beardedness, hirsuteness, muscularity, extraversion, and openness. The means and variance differ significantly between fathers and non-fathers in many other characteristics as well. This could be due to several reasons. First of all, it is possible that men with whom women reproduce actually differ from those with whom they do not. It should be noted, however, that most characteristics vary over time. This finding may thus be a side effect of higher age of fathers compared to non-fathers, especially in those characteristics where

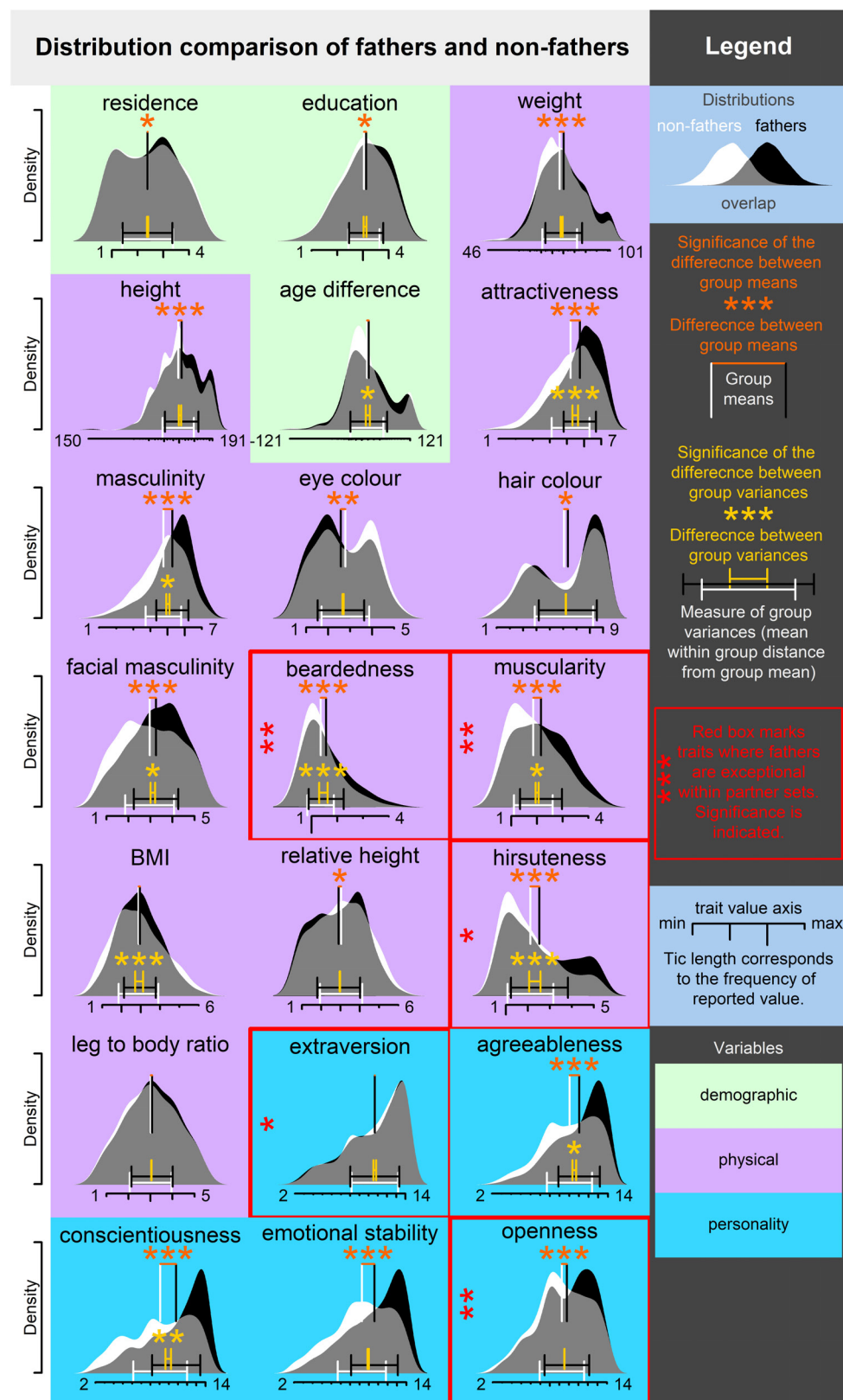


FIGURE 4 | Visualization of differences between fathers and non-fathers. Significance of difference between group means and variances is estimated from mixed effect models with respondent ID treated as a random factor. Significance levels are indicated as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 4 | Results of Mixed effect models comparing father/non-father means and variances.

	Comparison of mean values				Comparison of variances (within group residuals as in Levene's test)			
	Intercept non-father	Effect father	Standard error	p-value	Intercept non-father	Effect father	Standard error	p-value
Residence	2.48	−0.1	0.04	0.04	0.98	0	0.02	0.837
Education	3.03	0.1	0.04	0.015	0.61	0.05	0.03	0.064
Weight	78.18	1.98	0.45	<0.001	7.9	0.39	0.29	0.27
Height	179.66	1.19	0.31	<0.001	5.2	0.34	0.19	0.135
Age difference	34.52	3.67	2.31	0.132	32.97	3.99	1.47	0.02
Attractiveness	5.18	0.57	0.06	<0.001	1.1	−0.16	0.04	<0.001
Masculinity	4.73	0.55	0.06	<0.001	1.04	−0.1	0.04	0.025
Eye color	2.79	−0.21	0.06	0.001	1.08	−0.02	0.03	0.52
Hair color	6.05	0.3	0.12	0.018	2.19	−0.01	0.05	0.837
Facial masculinity	2.97	0.29	0.07	<0.001	1.11	−0.1	0.04	0.016
Beardedness	1.35	0.22	0.04	<0.001	0.52	0.17	0.02	<0.001
Muscularity	1.84	0.32	0.05	<0.001	0.75	0.06	0.03	0.046
BMI	2.88	0.07	0.06	0.238	1.02	−0.2	0.04	<0.001
Relative height	3.04	−0.14	0.06	0.025	1.08	0.03	0.04	0.52
Hirsuteness	2.12	0.41	0.07	<0.001	1.04	0.26	0.04	<0.001
Leg to body ratio	3	0.08	0.06	0.222	0.95	−0.01	0.04	0.837
Extraversion	10.67	0.01	0.16	0.971	2.5	0.17	0.08	0.086
Agreeableness	10.04	1	0.14	<0.001	2.32	−0.18	0.08	0.049
Conscientiousness	9.05	1.76	0.17	<0.001	2.94	−0.33	0.09	0.001
Emotional stability	9.22	1.39	0.16	<0.001	2.57	−0.08	0.09	0.514
Openness	9.23	0.56	0.14	<0.001	2.32	0.02	0.08	0.837

Respondent ID is treated as a random factor.

fathers disrupt mate choice consistency. Secondly, differences between fathers and non-fathers might be due to time-dependent cultural shifts. For instance fashions concerning beardedness vary significantly over time, which may cause a higher mutual similarity among former partners (non-fathers) as opposed to fathers (the most recent partner). Moreover, from an evolutionary perspective, these slight differences among partners could be due to the fact that each partner could be a potential father of a woman's children. It may be therefore beneficial for a woman not to experiment too much in her mate choice.

Nevertheless, some differences between fathers and non-fathers were observed. They could be due to individual relationship experience. In other words, it is possible that women adjust their mate choice depending on experiences gathered over lifetime and reproduce with a partner who fits their preferences better than earlier partners. Differences between fathers and non-fathers could also be due to memory bias or cognitive dissonance influenced by positive or negative experiences with particular partners. If so, the level of negative experiences with former partners should positively correlate with fathers' non-typicality. Moreover, women might have a tendency to ascribe more positive characteristics to a current partner (usually the father of her child or children) than to their ex-partners. In other words, partnership status itself may have an impact on the assessment. Alternatively, former partners could be regarded on average more positively simply because women's detailed memories of problems encountered in earlier relationships fade with time. There might be therefore some trade-offs between the principles

of 'my baby's father is always better' and 'sweet recollections of past loves.' We cannot address such possibilities in our analysis.

The fact that fathers lower the measured mate choice consistency and yet there is no meaningful systematic difference between fathers and non-fathers could be accounted for by either of two possible explanations. First of all, it is possible that women reproduce with men who have different characteristics than their ex-partners. This pattern was, however, found only for extroversion (whereby women who date extroverted men reproduced with more introverted individuals, while other women date introverts but reproduce with men who are more extroverted). Moreover, overall consistency of mate choice with respect to extroversion was high even when fathers were included in the partner sample and even in cases when fathers and non-fathers were excluded at random. It is then fair to assume that fathers do, after all, fit within the general type of women also in extraversion, although they tend to be on one of the extreme tails of this intrapersonal distribution. The second possible explanation is that variance in father and non-father group differs and fathers are a more variable group. We did not, however, encounter such a case in our dataset. If fathers had a higher variance than non-fathers, they would have to have also a higher average trait value. Where this was not the case (age difference), we found that father exclusion did significantly elevate mate choice consistency.

These findings are limited by including only women in reproductive age, because preferences and potentially also actual

choices can change in connection with changes in hormonal levels (e.g., menopause) during women's lives (Boothroyd and Vukovic, 2018). Female preferences are underpinned by a set of evolutionary adaptations (Kokko et al., 2003; Geary et al., 2004) and can change with age so as to reflect women's different interests (Kościński, 2011). Similarly, the importance of particular physical and personality characteristics can vary during one's life. To test intraindividual variability in mate choice, future studies should therefore investigate women's preferences and actual choices during their lives from childhood to menopause. Another limitation of our study is given by the fact that only respondents but not their (ex-)partners participated in the study. Although it would be nearly impossible to recruit also all (ex-)partners, it ought to be taken into account that when a woman reports about all of her partners during one session, this may lead to bias in a direction of mutual similarity. Alternatively, assessment of partners' characteristics could be biased by subsequent experiences, memories, or circumstances of a break-up of a relationship.

Our results support the hypothesis of consistency of mate choice with respect to a variety of characteristics, but further research is needed to confirm this effect through a longitudinal design. Secondly, consistency in mate choice should be investigated also in men and in a short-term mating context, where consistency of mate choice may be lower than in a long-term context (Štěrbová et al., 2018). Furthermore, future research should investigate interindividual differences in individual consistency (to find out which characteristics predict a consistent mate choice, the role of family members, etc.). In the light of all of the above, it would also be highly relevant to investigate to what degree are preferences inherited or learned. A twin study (Germine et al., 2015) has reported that face preferences seem to be mainly explained by environmental variation, but more research in this field is needed. And finally, research

should focus not only on actual choices but also on partner preferences.

AUTHOR CONTRIBUTIONS

ZŠ developed the study concept and collected data. PT performed the data analysis. ZŠ, PT, and KK interpreted results. ZŠ, PT, and KK drafted the manuscript. All authors approved the final version of the manuscript for submission.

FUNDING

This research was supported by the Czech Science Foundation project reg. no. 18-15168S, and by the Charles University Grant Agency project no. 1436317. This work was supported by Charles University Research Centre program no. 204056. ZŠ and PT were supported by the 'Sustainability for the National Institute of Mental Health' project, grant number LO1611, with financial support from the Ministry of Education, Youth, and Sports of the Czech Republic under the NPU I program.

ACKNOWLEDGMENTS

The authors wish to thank to Kristýna Taskovská for her help with data collection, Michaela Salák for assisting with participant recruitment, and Anna Pilátová for English proofreading.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00052/full#supplementary-material>

REFERENCES

- Allen, B., Curless, B., and Popović, Z. (2003). The space of human body shapes: reconstruction and parameterization from range scans. *ACM Trans. Graph.* 22, 587–594. doi: 10.1145/882262.882311
- Anderson, J. L., Crawford, C. B., Nadeau, J., and Lindberg, T. (1992). Was the duchess of windsor right? A cross-cultural review of the socioecology of ideals of female body shape. *Ethol. Sociobiol.* 13, 197–227. doi: 10.1016/0162-3095(92)90033-Z
- Bell, A. M., Hankison, S. J., and Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Anim. Behav.* 77, 771–783. doi: 10.1016/j.anbehav.2008.12.022
- Boothroyd, L. G., Jones, B. C., Burt, D. M., DeBruine, L. M., and Perrett, D. I. (2008). Facial correlates of sociosexuality. *Evol. Hum. Behav.* 29, 211–218. doi: 10.1016/j.evolhumbehav.2007.12.009
- Boothroyd, L. G., and Vukovic, J. (2018). "Mate preferences across the lifespan," in *The Oxford Handbook of Evolutionary Psychology and Behavioral Endocrinology*, eds L. W. Welling and T. K. Shackelford (Oxford: OUP).
- Botwin, M. D., Buss, D. M., and Shackelford, T. K. (1997). Personality and mate preferences: five factors in mate selection and marital satisfaction. *J. Pers.* 65, 107–136. doi: 10.1111/j.1467-6494.1997.tb00531.x
- Brumbaugh, C. C., and Wood, D. (2013). Mate preferences across life and across the world. *Soc. Psychol. Pers. Sci.* 4, 100–107. doi: 10.1177/1948550612442396
- Buss, D. M. (1989). Sex differences in human mate preferences: evolutionary hypotheses tested in 37 cultures. *Behav. Brain Sci.* 12, 1–14. doi: 10.1017/S0140525X00023992
- Buss, D. M., and Barnes, M. (1986). Preferences in human mate selection. *J. Pers. Soc. Psychol.* 50, 559–570. doi: 10.1037/0022-3514.50.3.559
- Buss, D. M. (1991). Evolutionary personality psychology. *Annu. Rev. Psychol.* 42, 459–491. doi: 10.1146/annurev.psych.42.1.459
- Christensen, T. L., Djurhuus, C. B., Clayton, P., and Christiansen, J. S. (2007). An evaluation of the relationship between adult height and health-related quality of life in the general UK population. *Clin. Endocrinol.* 67, 407–412. doi: 10.1111/j.1365-2265.2007.02901.x
- Conroy-Beam, D., Buss, D. M., Pham, M. N., and Shackelford, T. K. (2015). How sexually dimorphic are human mate preferences? *Pers. Soc. Psychol. Bull.* 41, 1082–1093. doi: 10.1177/0146167215590987
- Cotton, S., Small, J., and Pomiankowski, A. (2006). Sexual selection and condition-dependent mate preferences. *Curr. Biol.* 16, 755–765. doi: 10.1016/j.cub.2006.08.022
- Cunningham, M. R., Barbee, A. P., and Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial

- physical attractiveness. *J. Pers. Soc. Psychol.* 59, 61–72. doi: 10.1037/0022-3514.59.1.61
- Dixon, A. F., Halliwell, G., East, R., Wignarajah, P., Anderson, M. J., Alan, F., et al. (2003). Masculine somatotype and hirsuteness as determinants of sexual attractiveness to women. *Arch. Sex. Behav.* 32, 29–39. doi: 10.1023/A:1021889228469
- Dixon, B. J., and Brooks, R. C. (2013). The role of facial hair in women's perceptions of men's attractiveness, health, masculinity and parenting abilities. *Evol. Hum. Behav.* 34, 236–241. doi: 10.1016/j.evolhumbehav.2013.02.003
- Dixon, B. J., Dixon, A. F., Bishop, P. J., and Parish, A. (2010). Human physique and sexual attractiveness in men and women: a New Zealand-U.S. comparative study. *Arch. Sex. Behav.* 39, 798–806. doi: 10.1007/s10508-008-9441-y
- Dixon, B. J. W., and Rantala, M. J. (2015). The role of facial and body hair distribution in women's judgments of men's sexual attractiveness. *Arch. Sex. Behav.* 45, 877–889. doi: 10.1007/s10508-015-0588-z
- Eastwick, P. W., Harden, K. P., Shukusky, J. A., Morgan, T. A., and Joel, S. (2017). Consistency and inconsistency among romantic partners over time. *J. Pers. Soc. Psychol.* 112, 838–859. doi: 10.1037/pspi0000087
- Feingold, A. (1992). Good-looking people are not what we think. *Psychol. Bull.* 111, 304–341. doi: 10.1037/0033-2909.111.2.304
- Figueredo, A. J., Sefcek, J. A., and Jones, D. N. (2006). The ideal romantic partner personality. *Pers. Individ. Differ.* 41, 431–441. doi: 10.1016/j.paid.2006.02.004
- Frederick, D. A., and Haselton, M. G. (2007). Why is muscularity sexy? Tests of the fitness indicator hypothesis. *Pers. Social Psychol. Bull.* 33, 1167–1183. doi: 10.1177/0146167207303022
- Frost, P. (2006). European hair and eye color: a case of frequency-dependent sexual selection? *Evol. Hum. Behav.* 27, 85–103. doi: 10.1016/j.evolhumbehav.2005.07.002
- Frost, P., Kleisner, K., and Flegr, J. (2017). Health status by gender, hair color, and eye color: red-haired women are the most divergent. *PLoS One*, 12:e0190238. doi: 10.1371/journal.pone.0190238
- Geary, D. C., Vigil, J., and Byrd-Craven, J. (2004). Evolution of human mate choice. *J. Sex Res.* 41, 27–42. doi: 10.1080/00224490409552211
- Germin, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., et al. (2015). Individual aesthetic preferences for faces are shaped mostly by environments. *Not Genes. Curr. Biol.* 25, 2684–2689. doi: 10.1016/j.cub.2015.08.048
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the big-five personality domains. *J. Res. Pers.* 37, 504–528. doi: 10.1016/S0092-6566(03)00046-1
- Hensley, W. E. (1994). Height as a basis for interpersonal attraction. *Adolescence* 29, 469–474.
- Hönekopp, J. (2006). Once more: is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 199–209. doi: 10.1037/0096-1523.32.2.199
- Jennions, M. D., and Petrie, M. (2007). Variation in mate choice and mating preferences: a review of causes and consequences. *Biol. Rev.* 72, 283–327. doi: 10.1111/j.1469-185X.1997.tb00015.x
- Jensen-Campbell, L. A., and Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *J. Pers.* 69, 323–361. doi: 10.1111/1467-6494.00148
- Johnston, V. S. (2006). Mate choice decisions: the role of facial beauty. *Trends Cogn. Sci.* 10, 9–13. doi: 10.1016/j.tics.2005.11.003
- Judge, T. A., and Cable, D. M. (2004). The effect of physical height on workplace success and income: preliminary test of a theoretical model. *J. Appl. Psychol.* 89, 428–441. doi: 10.1037/0021-9010.89.3.428
- Kaplan, H. S., and Gangestad, S. W. (2005). "Life history theory and evolutionary psychology," in *The Handbook of Evolutionary Psychology*, ed. D. M. Buss (Hoboken, NJ: John Wiley, and Sons Inc), 68–95.
- Kenny, D. A., Kashy, D. A., and Cook, W. L. (2006). *Dyadic Data Analysis*. New York, NY: Guilford Press.
- Kleisner, K., Kočnar, T., Rubešová, A., and Flegr, J. (2010). Eye color predicts but does not directly influence perceived dominance in men. *Pers. Individ. Differ.* 49, 59–64. doi: 10.1016/j.paid.2010.03.011
- Kleisner, K., Priplatova, L., Frost, P., and Flegr, J. (2013). Trustworthy-looking face meets brown eyes. *PLoS One* 8:e53285. doi: 10.1371/journal.pone.0053285
- Kokko, H., Brooks, R., Jennions, M. D., and Morley, J. (2003). The evolution of mate choice and mating biases. *Proc. Biol. Sci.* 270, 653–664. doi: 10.1098/rspb.2002.2235
- Kościński, K. (2010). Do they know what they like? Intra-individual variation of female facial preferences. *J. Evol. Psychol.* 8, 23–55. doi: 10.1556/JEP.8.2010.1.4
- Kościński, K. (2011). Life history of female preferences for male faces. *Hum. Nat.* 22, 416–438. doi: 10.1007/s12110-011-9123-7
- Laeng, B., Mathisen, R., and Johnsen, J.-A. A. (2006). Why do blue-eyed men prefer women with the same eye color? *Behav. Ecol. Sociobiol.* 61, 371–384. doi: 10.1007/s00265-006-0266-1
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* 126, 390–423. doi: 10.1037/0033-2909.126.3.390
- Little, A. C. (2015). Attraction and human mating. in *Evolutionary Perspectives on Social Psychology*, eds V., Zeigler-Hill, L. L. M. Welling, and T. K. Shackelford (Cham: Springer), 319–332 doi: 10.1007/978-3-319-12697-5_25
- Newman, A. V., Pollet, T., McCarty, K., Neave, N., and Saxton, T. (2018). Isn't it eyeronic? Little evidence for consistent eye colour choices across relationships. *PsyArXiv [Preprint]*. doi: 10.31234/osf.io/dv93s
- Pawlowski, B. (2003). Variable preferences for sexual dimorphism in height as a strategy for increasing the pool of potential partners in humans. *Proc. R. Soc. B Biol. Sci.* 270, 709–712. doi: 10.1098/rspb.2002.2294
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887. doi: 10.1038/29772
- Pierce, C. A. (1996). Body height and romantic attraction: a meta-analytic test of the male-taller norm. *Soc. Behav. Pers. Int. J.* 24, 143–149. doi: 10.2224/sbp.1996.24.2.143
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Regan, P. C., Levin, L., Sprecher, S., Christopher, F. S., and Gate, R. (2000). Partner preferences. *J. Psychol. Hum. Sex.* 12, 1–21. doi: 10.1300/J056v12n03_01
- Rodriguez, R. L., Boughman, J. W., Gray, D. A., Hebets, E. A., Höbel, G., and Symes, L. B. (2013). Diversification under sexual selection: the relative roles of mate preference strength and the degree of divergence in mate preferences. *Ecol. Lett.* 16, 964–974. doi: 10.1111/ele.12142
- Rosenthal, G. G. (2017). *Mate Choice. The Evolution of Sexual Decision Making from Microbes to Humans*. New Jersey: Princeton University Press.
- Sarlio-Lahtenkorva, S., and Lahelma, E. (1999). The association of body mass index with social and economic disadvantage in women and men. *Int. J. Epidemiol.* 28, 445–449. doi: 10.1093/ije/28.3.445
- Saxton, T. K. (2016). Experiences during specific developmental stages influence face preferences. *Evol. Hum. Behav.* 37, 21–28. doi: 10.1016/j.evolhumbehav.2015.06.001
- Štěrbová, Z., Tureček, P., and Kleisner, K. (2018). Consistency of mate choice in eye and hair colour: testing possible mechanisms. *Evol. Hum. Behav.* 1, 74–81. doi: 10.1016/j.evolhumbehav.2018.08.003
- Štěrbová, Z., and Valentová, J. (2012). Influence of homogamy, complementarity, and sexual imprinting on mate choice. *Anthropologie* 50, 47–59.
- Swami, V., Einon, D., and Furnham, A. (2006). The leg-to-body ratio as a human aesthetic criterion. *Body Image* 3, 317–323. doi: 10.1016/j.bodyim.2006.08.003
- Tovée, M., Maisey, D., Vale, E., and Cornelissen, P. (1999). Characteristics of male attractiveness for women. *The Lancet* 353:1500. doi: 10.1016/S0140-6736(99)00438-9
- White, D., and Rabago-Smith, M. (2011). Genotype-phenotype associations and human eye color. *J. Hum. Genet.* 56, 5–7. doi: 10.1038/jhg.2010.126

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Štěrbová, Tureček and Kleisner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Do Single Men Smell and Look Different to Partnered Men?

Mehmet K. Mahmut* and Richard J. Stevenson

Food, Flavor and Fragrance Lab, Department of Psychology, Macquarie University, Sydney, NSW, Australia

Previous research indicates human body odor (BO) can signal kinship, sickness and genetic compatibility. Based on research indicating single males have higher testosterone levels than partnered males and that higher testosterone levels are associated with stronger smelling BO, the current study aimed to determine if, by extension of previous findings, single males' BO smells stronger than partnered males' BO. Eighty-two heterosexual women aged 18–35 years rated the BO and faces of six different males also aged 18–35 years. Consistent with the hypothesis, single men's BO smelled stronger than partnered men's BO and single men's faces were rated as more masculine than partnered men's faces. The possible advantages of females being able to identify single males are addressed in the Discussion.

Keywords: mate preferences, mate attraction, masculinity, body odor, face attractiveness

OPEN ACCESS

Edited by:

Kok Wei Tan,
University of Reading Malaysia,
Malaysia

Reviewed by:

Caroline Allen,
Newcastle University, United Kingdom
Jan Havlicek,
Charles University, Czechia
Shen Liu,
University of Science and Technology
of China, China

*Correspondence:

Mehmet K. Mahmut
mem.mahmut@mq.edu.au

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 September 2018

Accepted: 28 January 2019

Published: 13 February 2019

Citation:

Mahmut MK and Stevenson RJ
(2019) Do Single Men Smell and Look
Different to Partnered Men?
Front. Psychol. 10:261.
doi: 10.3389/fpsyg.2019.00261

INTRODUCTION

Humans rely heavily on visual cues to make mate preference judgements. From an evolutionary perspective, mate preferences based on facial attractiveness is advantageous for identifying and selecting a high quality partner (Buss and Schmidt, 1993). For example, research findings have demonstrated that facial attractiveness (Coetzee et al., 2009) and color (Stephen et al., 2011) are associated with physiological health. However, despite the vast majority of research focussing on signals detected by the visual sense, humans do not rely solely on visual cues to assess the suitability of a potential partner but also make judgements using their sense of smell (Stevenson, 2009). Specifically, the body odor (BO) of a potential partner is assessed by our sense of smell (Lübke and Pause, 2015) and given BOs can signal physical health and genetic compatibility with a potential partner, the role of BOs in mate attraction, and preference is not surprising.

In terms of our health, some infections (e.g., gangrene), and diseases (e.g., diabetic ketoacidosis) cause our bodies to emit odors that physicians can reliably recognize and use for diagnostic confirmation (Bijland et al., 2013). In terms of the genetic compatibility of a couple, a set of genes encoding the major histocompatibility complex (MHC) – cell-surface proteins involved in pathogen resistance (Milinski, 2006) that influence our BO (Milinski et al., 2013) – may also contribute to mate preference based on BO preference. For example, women have demonstrated a preference for the BO of men who have dissimilar MHC (Wedekind et al., 1996; Wedekind and Furi, 1997; Sorokowska et al., 2018) and offspring from MHC dissimilar (vs. similar) parents are potentially healthier. However, a recent meta-analysis (Winternitz et al., 2017) on the role of MHC in mate preference in various studies (not just those on BO-based preferences), concluded that mate choice was not driven by MHC differences.

Human BOs are not static and can change due to many factors, such as diet and menstrual cycle. For example, a study that experimentally controlled the amount of red meat consumed over a two-week period, found that a diet higher in meat is associated with unpleasant smelling

BO compared to a non-meat diet (Havlíček and Lenochova, 2006). However, it must be noted that Zuniga et al. (2017) found that higher meat consumption was associated with more pleasant smelling BO, although meat consumption frequency was based on self-report data which may account for the contrary findings to Havlíček and Lenochova (2006). Moreover, men's preferences for female BO vary based on the different stages of a women's menstrual cycle which are associated with the most dramatic changes in hormone levels; giving higher preference ratings for women's BO in the fertile phase of their cycle than those in the non-fertile phase (Gildersleeve et al., 2012).

While research investigating changes in hormone levels predominantly focus on the menstrual cycle, numerous studies have found differences in men's hormone levels based on their relationship status. Specifically, research findings have shown that heterosexual men with higher levels of testosterone were less likely to be married (Booth and Dabbs, 1993; Mazur and Michalek, 1998; van Anders and Watson, 2007; Van Anders and Goldey, 2010) or in long-term relationships (Gray et al., 2004) whereas lower levels of testosterone were associated with being in a romantic relationship. Further, various hormones (e.g., cortisol and testosterone) may affect the quality of a man's BO (Rantala et al., 2006) and how attractive they are perceived to be. For example, Thornhill et al. (2013) found that women's preference for BO of high testosterone men was significantly correlated ($r = 0.32$) with their probability of conception risk, presumably because higher testosterone may confer some form of evolutionary fitness (see Folstad and Karter, 1992). Similarly, Butovskaya et al. (2013) reported that women in the most fertile phase of their menstrual cycle prefer the BO of men with masculine qualities (e.g., social dominance) and numerous studies have shown women prefer BO of men with symmetrical faces (Gangestad and Thornhill, 1998; Thornhill and Gangestad, 1999; Thornhill et al., 2003).

In van Anders and Watson (2006) social neuroendocrinology theoretical framework, they presented evidence detailing the important role testosterone plays in behaviors that predict evolutionary fitness, namely; competition for resources, establishing a pair bond (securing a relationship), sexual activity plus parenting and pregnancy. A prediction arising from this conceptual framework is that higher testosterone levels are associated with competitive behaviors (such as acquiring resources) whereas lower testosterone levels are associated with pair-bond maintenance behaviors (such as intimate contact; van Anders and Watson, 2006). Given the evidence that men's hormone levels may differ based on their relationship status, and that hormone levels may in turn change the perceptual quality of men's BO, the aim of the current study was to empirically investigate for the first time whether single and partnered men's BO was perceptually different. Moreover, to assess the role that both visual and olfactory perception may play in mate preference, two modalities that are predominantly researched independently, the current study also tested whether the faces of single and partnered men differed based on visual ratings.

To determine whether single men's BO smelled different to the BO of partnered men, heterosexual female participants rated men's BO on five characteristics (e.g., sexiness, liking).

Based on previous research suggesting male testosterone levels were positively (but not significantly) associated with stronger smell BO ratings (Rantala et al., 2006) and single males have higher levels of testosterone (e.g., Booth and Dabbs, 1993), we hypothesized that single men's BO would smell stronger than that of partnered men's. Moreover, because stronger smelling BO ratings are associated with lower BO liking ratings (Havlíček and Lenochova, 2006), we predicted that single men's BO would be liked less and rated less sexy than partnered men's BO. In order to determine whether BO attractiveness predicted facial attractiveness, participants also rated the faces of the BO donors. Although the findings from three previous studies (Rikowski and Grammer, 1999; Thornhill and Gangestad, 1999; Foster, 2008) indicated the correlation between male BO and face attractiveness ratings made by fertile women is low (e.g., $r = 0.28$, $p = 0.030$; Thornhill and Gangestad, 1999), we hypothesized that favorable BO ratings (i.e., higher liking and sexiness) would be associated with favorable face ratings (e.g., attractive, masculine). We made no *a priori* predictions about differences between single and partnered men's face attractiveness ratings. Finally, to ensure the ability to compare the BO and face ratings of single and partnered men, participants rated the stimuli of three single and three partnered unknown men.

MATERIALS AND METHODS

Participants

Eight-two (42 single, 40 partnered) heterosexual females ($M = 20.2$ years, $SD = 2.9$) completed the study at Macquarie University for credit towards an introductory psychology course. A single participant was someone who was not in a committed romantic relationship whereas a partnered participant was someone who was in a monogamous, romantic relationship. Given single and partnered women may perceive a man's BO or face differently (e.g., Little et al., 2002) we included both partnered and single women in this study. Participants were asked about their medical history and to indicate whether their sense of smell functioned normally. Only heterosexual females aged between 18 and 35 years, who indicated they had a normal sense of smell with no history of a condition, injury or surgery that compromised their sense of smell prior to, or on the day of the study, qualified for the study. Clearance to conduct the study was granted by the Human Research Ethics Committee at Macquarie University's and all participants and donors gave written and informed consent.

Donors of Body Odor and Face Pictures

The BOs and face pictures of 91 males formed the stimuli pool for the current study. The donors had no other involvement in the study aside from supplying their BO and face picture. The majority of donors were selected by participants; for partnered participants, the donor was their current partner and for single participants, the donor was their friend or brother. However, the Experimenters also recruited 10 donors to ensure there was a sufficiently large stimulus pool to draw from. All donors had to be aged between 18 and 35 years to qualify for the study. All donors

were heterosexual, except for one who identified as homosexual, whose BO was included in the stimulus pool. Overall, 46 of the BO donors were single and 45 were partnered. However, there was no significant difference between single and partnered donors in terms of their Body Mass Index (BMI; 24.8 vs. 24.3) or age (21 vs. 22.5 years).

Donor Data, Stimuli Collection and Preparation

Body Odor Collection and Preparation

Approximately one week before testing, each participant collected a donor pack from the Experimenter. The donor pack included a new, white, 100% cotton T-shirt in a resealable plastic bag, an instructions sheet and short survey containing demographic questions which participants delivered it to their known donor. Odor donors were instructed to avoid eating odorous foods (e.g., garlic, onion; Fialová et al., 2016) 24 h before and while wearing the T-shirt, wash using non-perfumed products before wearing the T-shirt and not to use perfumed products while wearing the T-shirt (Allen et al., 2016). The donor was instructed to wear the T-shirt for one day (i.e., no more than 24 h) and to not remove the shirt until a significant amount of sweat was absorbed onto the underarm of the T-shirt. The instruction sheet included a photograph of a model wearing a white T-shirt depicting an unacceptable amount of underarm sweat (i.e., approximately 25% of underarm patch appeared wet with sweat) and the minimum acceptable amount of underarm sweat (i.e., approximately 75% of underarm patch appeared wet with sweat). The type of physical activity participants engaged in to produce the sweat was not prescribed but it was suggested that brisk walking or sporting activities may facilitate sweating.

After removing the T-shirt, donors were asked to return the T-shirt to the resealable plastic bag provided and immediately store in a freezer. Participants collected the sweated-in T-shirt from donors and brought it in on the day of testing. Participants were informed of the importance of keeping the shirt in a freezer until bringing it into the lab. Upon receiving the T-shirt, the Experimenter cut out both underarms of the T-shirt and placed each in a new separate, opaque, plastic condiment bottle that was approximately 14 cm tall with a 250 mL capacity. Each bottle had a screw-on lid with an elongated nozzle with a removal cap and a 5mm opening through which the odorant was delivered. When not in use, the bottles stored in a freezer, a method validated in previous studies (e.g., Lenochova et al., 2009).

Face Pictures

Donors also supplied a current, digital, color, passport-style photo (i.e., neutral face, no hat or glasses) which was digitally adjusted using a computer to a height of 8 cm before being printed (in color) on white, A4-sized paper.

Donor Demographics

Each donors' height, weight, age, relationship status (i.e., single or partnered) and relationship to participant (i.e., partner, friend, relative) was collected via a short self-report survey that was included in the donor pack.

Measures

Excluded Participants and Variables

Two partnered participants' data were excluded from analyses because one's partner was not within the accepted age range (of 18 to 35 years) and the other returned a T-shirt smelling of perfume. Other measures were administered as part of a larger project, namely self-report measures relating to the nature of the donor-target relationship. The results from these measures were unrelated to the aims and hypotheses of the current study and are therefore not reported here. Finally, to remove any bias associated with preference a participant may have for their donor's BO and/or face, the results presented do not include the data from the ratings participants made of their donor.

Body Odor and Face Stimuli Selection

The Experimenter selected six different donors' BO and the six corresponding face pictures which each Participant would be presented in a random order. The first BO and face picture selected belonged to the participant's donor. The BOs and faces of the next six donors (three single, three partnered) were randomly selected from two separate donor pools; one consisting of single and the other consisting of partnered donors unknown to the participant.

Body Odor Characteristics Ratings Task

The six BOs were randomly presented to participants who made five ratings of each BO based on the following questions (variable label in brackets): (1) How much do you like/dislike this smell? ("Like"); (2) How sexy does this odor smell? ("Sexy"); (3) How familiar are you with this smell? ("Familiarity"); (4) How strong does this smell? ("Strong") (5) How much does this odor smell like your odor donor? ("Similarity"), on a 7-point scale from zero (not at all) to six (very). The Experimenter squeezed the bottle containing the BO three times approximately 2.5 cm from participants' nostrils while participants inhaled through their nose. The minimum inter-stimulus interval was 30-s. For each of the five BO characteristics ratings, two variables were computed: the first was the average rating given by the participant to the BO of partnered donors and the second was the average rating given by the participant to the BO of single donors. Therefore, a total of 10 variables were computed. For example, for the BO "Like" ratings, there were two variables created: one was the BO "Like" rating averaged across all single donors that were rated and the second variable created was the BO "Like" rating averaged across all partnered donors that were rated.

Face Characteristics Ratings Task

Participants were randomly presented with the six faces corresponding to the six BOs selected and asked to rate each face on eight characteristics that have been found to be universally desired (Buss, 1989, 1994; i.e., Masculine, Good Partner, Sexy, Intelligent, Loyal, Attractive, Kind and Trustworthy) on a scale ranging from zero (not at all) to six (very). For each of the eight face characteristics ratings, two variables were computed: the first was the average rating given by the participant to the faces of partnered donors and the second was the average rating given

by the participant to the faces of single donors. Therefore, a total of 16 variables were computed. For example, for the face Masculine ratings, there were two variables created: one was the face Masculine rating averaged across all single donors that were rated and the other was the face Masculine rating averaged across all partnered donors that were rated.

Procedure

The study was administered by three different female Experimenters, each conducting a similar number of studies.

Preliminary Data Analysis

Note that we tested whether having a beard influenced face masculinity ratings by comparing face masculinity scores of donors with beards (9% of sample) and without (91% of the sample); the results of an independent samples *t*-tests revealed no significant differences between these groups (all *ps* > 0.05). We also tested, but found no significant differences, between partnered and single female participants or between females using or not using birth contraception in terms of their ratings of single and partnered men's BO and faces.

RESULTS

Were Single and Partnered Men's BO Rated Differently by Single and Partnered Women?

To determine whether single and partnered female participants rated single and partnered men's BO differently on five characteristics (i.e., Strong, Like, Sexy, Familiarity and Similarity), five 2×2 mixed design analysis of variances (ANOVA) were ran (see **Table 1** for descriptive statistics). The between-subjects variable in each ANOVA was Participant Relationship Status (i.e., partnered or single) and the within-subjects variable was Donor Relationship Status which

had two levels (i.e., partnered or single). The family-wise error rate was adjusted for the five comparisons made such that the alpha-level was set at 0.01 (i.e., 0.05/5).

The first ANOVA was conducted with the BO strong ratings as the dependent variable (DV), which revealed a significant main effect for Donor Relationship Status, $F(1,77) = 9.51$, $p = 0.003$, $\eta_p^2 = 0.11$, indicating that averaged across participants, single men's BO was rated as smelling stronger than partnered donor's BO. There was no significant main effect for Participant Relationship Status, $F(1,77) = 2.16$, $p = 0.15$, $\eta_p^2 = 0.03$, or Participant Relationship Status \times Donor Relationship Status interaction ($F < 1$).

The next four ANOVAs revealed no significant main or interaction effects (with 11 of 12 *F*-values < 1) indicating that partnered and single women did not rate partnered and single men's BO different on BO characteristic ratings of Like, Sexy, Familiarity and Similarity.

Were Single and Partnered Men's Faces Rated Differently by Single and Partnered Women?

To determine whether single and partnered female participants rated single and partnered men's faces differently on eight characteristics (i.e., Masculine, Good Partner, Sexy, Intelligent, Loyal, Attractive, Kind, And Trustworthy), eight separate 2×2 mixed design analysis of variances (ANOVA) were conducted (see **Table 2** for descriptive statistics). The between-subjects variable in each ANOVA was Participant Relationship Status (i.e., partnered or single) and the within-subjects variable was Donor Relationship Status (i.e., partnered or single). The family-wise error rate was adjusted for the eight comparisons made such that the alpha-level was set at 0.006 (i.e., 0.05/8).

The first ANOVA was conducted with the face Masculine ratings as the DV, revealing a significant main effect for Donor Relationship Status, $F(1,77) = 18.76$, $p < 0.001$, $\eta_p^2 = 0.20$

TABLE 1 | Single and partnered women's ratings of single and partnered men's body odor.

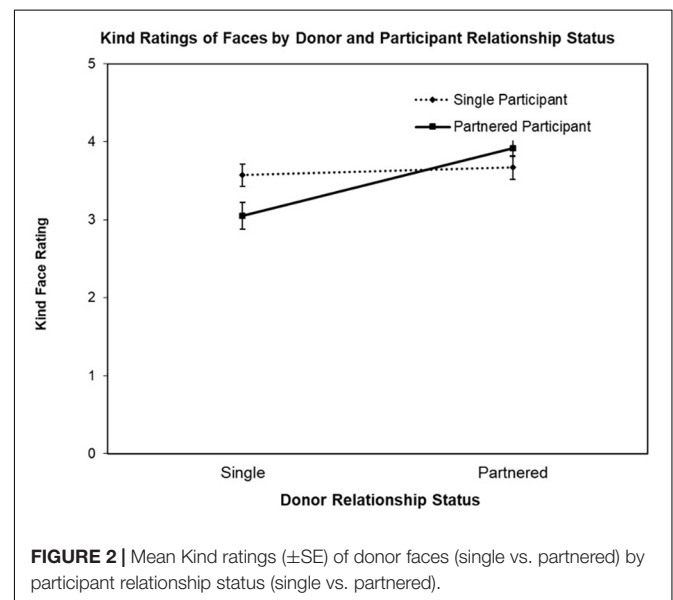
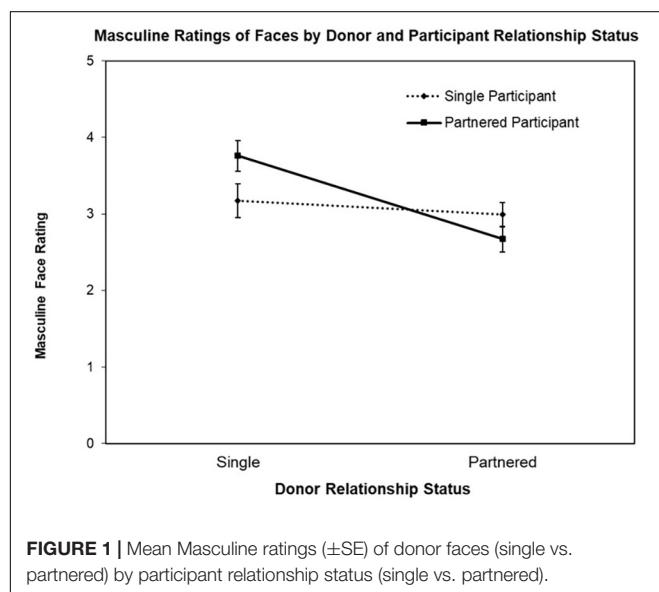
BO ratings	Entire sample (<i>n</i> = 82)		Partnered women (<i>n</i> = 40)		Single women (<i>n</i> = 42)	
	Donor type (Male)		Donor type (Male)		Donor type (Male)	
	Single	Partnered	Single	Partnered	Single	Partnered
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Strong	3.54 (1.14)	3.04 (1.06)	3.64 (1.12)	3.21 (1.07)	3.43 (1.16)	2.86 (1.05)
Familiarity	1.76 (1.36)	1.76 (1.02)	1.61 (1.42)	1.69 (1.10)	1.93 (1.29)	1.83 (0.94)
Sexy	1.41 (1.20)	1.44 (1.14)	1.37 (1.20)	1.37 (1.14)	1.46 (1.22)	1.51 (1.14)
Like	2.08 (1.27)	2.21 (1.02)	2.04 (1.31)	2.08 (1.07)	2.13 (1.24)	2.34 (0.97)
Similarity	1.61 (1.30)	1.65 (1.08)	1.49 (1.37)	1.76 (1.17)	1.73 (1.22)	1.53 (0.98)

TABLE 2 | Single and partnered women's ratings of single and partnered men's faces.

Face ratings	Entire sample (Women) (<i>n</i> = 82)		Partnered women (<i>n</i> = 40)		Single women (<i>n</i> = 42)	
	Donor relationship status (Male)		Donor relationship status (Male)		Donor relationship status (Male)	
	Single	Partnered	Single	Partnered	Single	Partnered
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Masculine	3.47 (1.25)	2.83 (1.05)	3.76 (1.27)	2.67 (1.07)	3.17 (1.17)	2.99 (1.02)
Good partner	1.71 (1.18)	1.80 (1.23)	1.59 (1.24)	1.50 (1.16)	1.83 (1.11)	2.10 (1.02)
Sexy	2.11 (1.34)	1.80 (1.08)	2.18 (1.36)	1.72 (0.99)	2.04 (1.33)	1.89 (1.16)
Intelligent	3.83 (0.80)	3.77 (0.81)	3.84 (0.77)	3.92 (0.74)	3.83 (0.84)	3.62 (0.87)
Loyal	2.94 (1.07)	3.18 (1.02)	2.91 (0.93)	3.44 (0.56)	2.97 (1.20)	2.94 (1.28)
Attractive	2.42 (1.27)	2.28 (1.15)	2.41 (1.35)	2.23 (1.12)	2.43 (1.19)	2.34 (1.89)
Kind	3.31 (1.03)	3.79 (0.80)	3.05 (1.10)	3.92 (0.67)	3.57 (0.89)	3.67 (0.90)
Trustworthy	3.15 (1.02)	3.41 (0.86)	3.02 (1.10)	3.59 (0.80)	3.27 (0.91)	3.22 (0.88)

and interaction for Donor Relationship Status by Participant Relationship Status, $F(1,77) = 9.70$, $p = 0.003$, $\eta_p^2 = 0.11$ (see **Figure 1**). The main effect for Participant Relationship Status was not significant ($F < 1$). Follow-up contrast testing revealed the nature of the interaction, that is, while partnered female participants rated single men's faces as more masculine than partnered men's faces, $t(39) = 5.72$, $p < 0.001$, $d' = 0.93$, single women did not rate partnered and single men's faces differently on Masculine, $t < 1$.

The second ANOVA was conducted with the face Kind ratings as the DV, revealing a significant main effect for Donor Relationship Status, $F(1,77) = 14.95$, $p < 0.003$, $\eta_p^2 = 0.16$ and interaction for Donor Relationship Status by Participant Relationship Status, $F(1,77) = 9.53$, $p = 0.003$, $\eta_p^2 = 0.11$ (see **Figure 2**). The main effect for Participant Relationship Status was not significant ($F < 1$). Follow-up contrast testing revealed the nature of the interaction, that is, while partnered female participants rated partnered men's faces as appearing kinder



than single men's faces, $t(39) = 4.94$, $p < 0.001$, $d' = 0.95$, single women did not rate partnered and single men's faces differently on Kind, $t < 1$.

The third ANOVA was conducted with the face Trustworthy ratings as the DV, revealing that the main effect for Donor Relationship Status was not significant, $F(1,77) = 3.68$, $p = 0.059$, $\eta_p^2 = 0.0$, nor was the main effect for Participant Relationship Status ($F < 1$). While a significant interaction for Donor Relationship Status by Participant Relationship Status was found [$F(1,77) = 5.45$, $p = 0.022$, $\eta_p^2 = 0.07$], this effect was not significant based on the adjusted alpha-level.

The fourth ANOVA was conducted with the face Loyalty ratings as the DV, revealing that the main effect for Donor Relationship Status was not significant, $F(1,77) = 3.87$, $p = 0.053$, $\eta_p^2 = 0.05$, nor was the Participant Relationship Status main effect, $F(1,77) = 1.30$, $p = 0.26$, $\eta_p^2 = 0.02$. While a significant interaction for Donor Relationship Status by Participant Relationship Status was found [$F(1,77) = 4.59$, $p = 0.035$, $\eta_p^2 = 0.05$], this interaction effect was rendered non-significant based on the adjusted alpha-level.

The next four ANOVAs conducted revealed no significant main or interaction effects (with 8 of 12 F -values < 1) indicating that partnered and single women did not rate partnered and single men's faces different on ratings of good partner, sexy, intelligent, and attractive.

Exploratory Analyses: Do BO Ratings Predict Face Ratings?

In order to determine whether BO Characteristics ratings predicted Face Characteristics ratings, a Spearman's rho correlation analysis was conducted, which overall, revealed favorable BO ratings (i.e., Sexy and Like) were associated with favorable face ratings (e.g., Attractive, Intelligent; see Table 3). For example, higher BO Like ratings were significantly correlated with rating faces more Attractive ($r = 0.29$, $p = 0.008$), Masculine ($r = 0.30$, $p = 0.007$), Sexy ($r = 0.26$, $p = 0.019$), and someone who would make a Good Partner ($r = 0.33$, $p = 0.003$). The

inter-correlations among the face ratings were positive and statistically significant (except for four); the lowest was between Intelligent and Good Partner ($r = 0.12$, $p = 0.30$) and the highest was between Sexy and Attractive ($r = 0.89$, $p < 0.001$). The inter-correlations among the BO ratings were mostly positive and statistically significant, except for those with the Strong ratings. The lowest significant correlation was between Familiarity and Sexy ratings ($r = 0.48$, $p < 0.001$) and the highest was between Sexy and Like ($r = 0.78$, $p < 0.001$; see Table 3).

The ANOVA results reported above demonstrated that partnered and single participants rated partnered and single donors differently, specifically on the BO Strong ratings and a subset of the face ratings (i.e., Masculine, Loyal, Kind, and Trustworthy). Therefore, we explored the correlations amongst the ratings indicated by the ANOVA findings to determine the nature of the differences between partnered and single females' ratings. This exploration revealed that the largest discrepancies were all based on ratings of partnered donors' BO and faces. The largest discrepancy was the correlation between BO Strong and Face Trustworthy ratings: specifically, for the ratings given by partnered women, we found a negative correlation ($r = -0.35$, $p = 0.025$) whereas for the ratings given by single women, we found a positive correlation ($r = 0.11$, $p = 0.51$). A Fisher's r -to- z transformation comparison test indicated these two correlations were significantly different ($Z = 2.07$), confirming that higher BO Strong ratings were associated with lower Face Trustworthy ratings for partnered women but no such relationship existed for single women. While there were other similarly large discrepancies between partnered and single participants' ratings, none were significantly different.

DISCUSSION

Consistent with our hypothesis, single men's BO was rated as smelling stronger than the BO of partnered men. We also found that single men's faces were rated as more masculine

TABLE 3 | Body odor and face ratings correlations ($N = 82$).

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
(1) Sexy BO	0.48**	0.78**	0.50**	-0.02	0.25*	0.33**	0.30**	0.17	0.06	0.31**	0.30**	0.08
(2) Familiarity BO		0.52**	0.72**	0.06	-0.02	0.11	0.06	0.10	-0.05	0.18	-0.04	0.09
(3) Like BO			0.56**	-0.10	0.29**	0.33**	0.24*	0.15	0.06	0.30**	0.26*	0
(4) Similarity BO				0.12	0.15	0.34**	0.12	0.16	0.08	0.46**	0.10	0.10
(5) Strong BO					0.15	0.09	0.16	0.13	0.11	0.17	0.14	0.12
(6) Attractive Face						0.70**	0.27*	0.32**	0.31**	0.48**	0.89**	0.19
(7) Good Partner Face							0.23*	0.25*	0.23*	0.47**	0.68**	0.12
(8) Intelligent Face								0.57**	0.53**	0.22*	0.29**	0.55**
(9) Kind Face									0.66**	0.31**	0.32**	0.82**
(10) Loyal Face										0.28*	0.30**	0.71**
(11) Masculine Face											0.44**	0.20
(12) Sexy Face												0.19
(13) Trustworthy Face												

Results based on Spearman's rho correlations. * $p < 0.05$, ** $p < 0.01$. BO = body odor.

than partnered men's faces, but only among partnered women. Moreover, partnered women rated partnered men's faces as kinder, more trustworthy and loyal than single men's faces, but single females rated partnered and single men's faces similarly on these characteristics. Finally, the results showed favorable BO ratings were correlated with favorable ratings of the corresponding faces. Although testosterone levels were not directly tested here, the current study's findings are congruent with previous research showing that single and partnered males can be differentiated based on their testosterone levels (e.g., Van Anders and Goldey, 2010), that higher testosterone levels are associated with a stronger smelling BO (Rantala et al., 2006) and that more intense BOs are rated more masculine smelling (Havlíček and Lenochova, 2006).

An obvious question is; why would a single male's BO smell different from that of a partnered man's BO? The social neuroendocrinology theoretical framework (van Anders and Watson, 2006) helps frame a possible answer to this question. Specifically, BOs are the manifestation of our current endocrinology (e.g., low or high testosterone levels) which signal the fitness, viability, and/or availability of a potential mate. Based on their study's results, Van Anders and Goldey (2010) concluded that single males have higher levels of testosterone than partnered males because of the sexual competition associated with being single and that low testosterone levels are associated with bond maintenance. From an evolutionary perspective, it may be advantageous for women to be able to detect the chemosignals that connote coupledness and ultimately avoid courting partnered males (especially with offspring) due to the relatively reduced resources they can offer.

An alternative explanation is that single men's BO may smell more intense than partnered men's BO because of their poorer health and/or hygiene. Evidence for this assertion comes from research showing single men have poorer physical and mental health outcomes than partnered men (Hu and Goldman, 1990) which may manifest as poorer hygiene and therefore BO. Further evidence comes from research showing married men are significantly more likely to seek health care due to their wives' influence compared to unmarried men (Norcross et al., 1996). While we found no evidence that single men were less healthy than partnered men based on the fact there were no significant group differences in terms of BMI, the positive health impact of having a partner may explain our findings.

The current study's finding that single men's faces were rated significantly more masculine than partnered men's faces (among partnered women only) is consistent with previous research showing higher testosterone levels are associated with more intense smelling BO (Rantala et al., 2006); especially when considered in conjunction with the finding that single men have higher levels of testosterone than partnered men (e.g., Van Anders and Goldey, 2010). Given higher testosterone levels are associated with more masculine facial features (Penton-Voak and Chen, 2004), it is possible single males in the current sample had higher levels of testosterone. However, a single man's facial features are unlikely to change overnight unlike their relationship

status, so alternative explanations for the differences in Masculine ratings for partnered and single men's faces must be considered. While facial features do change with age, partnered males were not older than single males so age can be ruled out as an explanation for group differences in facial masculinity. Having a beard was also excluded as an explanation for higher Masculine ratings of single men's faces but it remains possible that individual differences in what constitutes a "masculine" face may, to some extent, account for the finding.

While it is curious that only partnered women rated single men's faces as more Masculine than partnered men's faces, previous research indicates partnered women in the fertile phase of their menstrual cycle (compared to those in their non-fertile phase) find single men's faces more attractive than partnered men's faces, especially if they are masculine-versus feminine-looking (Bressan and Stranieri, 2008). A limitation of the current study was that participants' menstrual cycles were not assessed so we cannot conclude whether menstrual cycle phase influenced their face masculinity ratings. Further limitations include not supplying donors with non-perfumed body cleansing products or specifying a specific duration of exercise, which may have contributed to variability in the quality and nature of the stimuli collected.

The correlations between BO and face ratings revealed a consistent pattern of results indicating favorable BO ratings were associated with favorable face ratings. While the current study's findings are congruent with previous findings, the positive relationship between BO and face ratings has largely been demonstrated with female participants in the fertile phase of their menstrual cycle (Rikowski and Grammer, 1999; Thornhill and Gangestad, 1999). However, the correlation between BO like ratings and face attractiveness ratings for low fertile compared to high fertile women in both studies (i.e., Rikowski and Grammer, 1999; Thornhill and Gangestad, 1999) were not significantly different, suggesting no reliable group differences. Moreover, Allen et al. (2016) found women's ratings of masculinity for men's BO was positively and significantly correlated with face masculinity ratings, although the women's menstrual phase was not recorded in their study, either. As we could not compare the BO and face rating correlations based on a participant's menstrual phase, it remains possible that differences exist between the low and high fertility phases of the menstrual cycle.

While the current results show a single man's BO smells more intense and their face appears more masculine than a partnered man's, the findings are preliminary and require replication. A specific aim of future research would be to determine whether testosterone levels are responsible for the differences in BO and face ratings between single and partnered men found in the current study. This could be achieved in a single study using the same participants with the aim to (a) replicate the finding that single men's BO smells more intense than partnered men's BO; (b) replicate the finding that single men's faces are rated more masculine than partnered men's faces; (c) confirm that single men have higher levels of testosterone than partnered men; (d) assess women's menstrual

cycle phase, and (e) comparing an individual's BO while single and coupled. Future studies would also benefit from ruling out alternative explanations for BO differences between single and partnered men, such as those associated with poor physical and mental health.

AUTHOR CONTRIBUTIONS

MM and RS was involved in the study design, data and analyses, and production and editing of the final document.

REFERENCES

- Allen, C., Cobey, K. D., Havlíček, J., and Roberts, S. C. (2016). The impact of artificial fragrances on the assessment of mate quality cues in body odor. *Evol. Hum. Behav.* 37, 481–489. doi: 10.1016/j.evolhumbehav.2016.05.001
- Bijland, L. R., Bomers, M. K., and Smulders, Y. M. (2013). Smelling the diagnosis: a review of scent in diagnosing disease. *Neth. J. Med.* 71, 300–306.
- Booth, A., and Dabbs, J. (1993). Testosterone and men's marriages. *Soc. Forces* 72, 463–477. doi: 10.1093/sf/72.2.463
- Bressan, P., and Stranieri, D. (2008). The best men are (not always) already taken: female preference for single versus attached males depends on conception risk. *Psychol. Sci.* 19, 145–151. doi: 10.1111/j.1467-9280.2008.02060.x
- Buss, D. M. (1989). Sex differences in human mate preference: evolutionary hypotheses tested in 37 cultures. *Behav. Brain Sci.* 12, 1–14. doi: 10.1017/S0140525X00023992
- Buss, D. M. (1994). "Mate preferences in 37 cultures," in *Psychology and Culture*, eds R. S. Malpass and W. J. Lonner (Boston: Allyn and Bacon), 197–201.
- Buss, D. M., and Schmidt, D. P. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychol. Rev.* 100, 204–232. doi: 10.1037/0033-295X.100.2.204
- Butovskaya, M. L., Veselovskaya, E. V., Rostovtseva, V. V., Selverova, N. B., and Ermakova, I. V. (2013). Mechanisms of human reproductive behavior: olfactory markers of male attractiveness. *Biol. Bull. Rev.* 33, 196–208. doi: 10.1134/S2079086413030031
- Coetsee, V., Perrett, D. I., and Stephen, I. D. (2009). Facial adiposity: a health cue? *Perception* 38, 1700–1711.
- Fialová, J., Roberts, S. C., and Havlíček, J. (2016). Consumption of garlic positively affects hedonic perception of axillary body odour. *Appetite* 97, 8–15. doi: 10.1016/j.appet.2015.11.001
- Folstad, I., and Karter, A. J. (1992). Parasites, bright males and the immunocompetence handicap. *Am. Nat.* 139, 603–622. doi: 10.1086/285346
- Foster, J. D. (2008). Beauty in mostly in the eye of the beholder: olfactory versus visual cues of attractiveness. *J. Soc. Psychol.* 148, 765–773. doi: 10.3200/SOCP.148.6.765-774
- Gangestad, S. W., and Thornhill, R. (1998). Menstrual cycle variation in women's preference for the scent of symmetrical men. *Proc. Biol. Sci.* 265, 927–933. doi: 10.1098/rspb.1998.0380
- Gildersleeve, K. A., Haselton, M. G., Larson, C. M., and Pillsworth, E. G. (2012). Body odor attractiveness as a cue for impending ovulation in women: evidence from a study using hormone-confirmed ovulation. *Horm. Behav.* 61, 157–166. doi: 10.1016/j.yhbeh.2011.11.005
- Gray, P. B., Campbell, B. C., Marlowe, F. W., Lipson, S. F., and Ellison, P. T. (2004). Social variables predict between-subject but not day-to-day variation in the testosterone of US men. *Psychoneuroendocrinology* 29, 1153–1162. doi: 10.1016/j.psyneuen.2004.01.008
- Havlíček, J., and Lenochova, P. (2006). The effect of meat consumption on body odor attractiveness. *Chem. Senses* 31, 747–752. doi: 10.1093/chemse/bjl017
- Hu, Y., and Goldman, N. (1990). Mortality differentials by marital status: an international comparison. *Demography* 27, 233–250. doi: 10.2307/2061451
- Lenochova, P., Roberts, S. C., and Havlíček, J. (2009). Methods of human body odor sampling: the effects of freezing. *Chem. Senses* 34, 127–138. doi: 10.1093/chemse/bjn067
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. Biol. Sci.* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Lübke, K. T., and Pause, B. M. (2015). Always follow your nose: the functional significance of social chemosignals in human reproduction and survival. *Horm. Behav.* 68, 134–144. doi: 10.1016/j.yhbeh.2014.10.001
- Mazur, A., and Michalek, J. (1998). Marriage, divorce and male testosterone. *Soc. Forces* 77, 315–330. doi: 10.1093/sf/77.1.315
- Milinski, M. (2006). The major histocompatibility complex, sexual selection and mate choice. *Annu. Rev. Ecol. Syst.* 37, 159–186. doi: 10.1146/annurev.ecolsys.37.091305.110242
- Milinski, M., Croy, I., Hummel, T., and Boehm, T. (2013). Major histocompatibility complex peptide ligands as olfactory cues in human body odour assessment. *Proc. Biol. Sci.* 280:2889. doi: 10.1098/rspb.2012.2889
- Norcross, W. A., Ramirez, C., and Palikas, L. A. (1996). The influence of women on the health care-seeking behavior of men. *J. Family Pract.* 43, 475–480.
- Penton-Voak, I. S., and Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evol. Hum. Behav.* 25, 229–241. doi: 10.1016/j.evolhumbehav.2004.04.003
- Rantala, M. J., Eriksson, C. J. P., Vainikka, A., and Kortet, R. (2006). Male steroid hormones and female preference for male body odor. *Evol. Hum. Behav.* 27, 259–269. doi: 10.1016/j.evolhumbehav.2005.11.002
- Rikowski, A., and Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proc. Biol. Sci.* 266, 869–874. doi: 10.1098/rspb.1999.0717
- Sorokowska, A., Pietrowski, D., Schäfer, L., Kromer, J., Schmidt, A. H., Sauter, J., et al. (2018). Human leukocyte antigen similarity decreases partners' and strangers' body odor attractiveness for women not using hormonal contraception. *Horm. Behav.* 106, 144–149. doi: 10.1016/j.yhbeh.2018.10.007
- Stephen, I. D., Coetsee, V., and Perrett, D. I. (2011). Carotenoid and melanin pigment coloration affect perceived human health. *Evol. Hum. Behav.* 32, 216–227. doi: 10.1016/j.evolhumbehav.2010.09.003
- Stevenson, R. J. (2009). An initial evaluation of the functions of human olfaction. *Chem. Senses* 35, 3–20. doi: 10.1093/chemse/bjp083
- Thornhill, R., Chapman, J. F., and Gangestad, S. W. (2013). Women's preferences for men's scents associated with testosterone and cortisol levels: patterns across the ovulatory cycle. *Evol. Hum. Behav.* 34, 216–221. doi: 10.1016/j.evolhumbehav.2013.01.003
- Thornhill, R., and Gangestad, S. W. (1999). The scent of symmetry: a human sex pheromone that signals fitness? *Evol. Hum. Behav.* 20, 175–201. doi: 10.1016/S1090-5138(99)00005-7
- Thornhill, R., Gangestad, S. W., Miller, R., Scheyd, G., Knight, J., and Franklin, M. (2003). MHC, symmetry, and body scent attractiveness in men and women. *Behav. Ecol.* 14, 668–678. doi: 10.1093/beheco/arg043
- Van Anders, S. M., and Goldey, K. L. (2010). Testosterone and partnering are linked via relationship status for women and 'relationship orientation' for me. *Horm. Behav.* 58, 820–826. doi: 10.1016/j.yhbeh.2010.08.005

ACKNOWLEDGMENTS

We thank Macquarie University for providing funding to conduct the current study (grant no. 9201400673). Macquarie University had no involvement in the study design, data collection, data analysis, data interpretation, writing of this manuscript, or selection of the journal to submit the manuscript to. We also thank the meticulous, detailed and constructive feedback received from the Reviewers. Finally, we would like to thank the Research Assistants who ran the study; Dr. Samantha Adams, Dervisen Komuksu, and Madeleine Fraser.

- van Anders, S. M., and Watson, N. V. (2006). Social neuroendocrinology: effects of social behaviors and contexts on sex steroids in humans. *Hum. Nat.* 17, 212–237. doi: 10.1007/s12110-006-1018-7
- van Anders, S. M., and Watson, N. V. (2007). Testosterone levels in women and men who are single, in long-distance relations, or same-city relationships. *Horm. Behav.* 51, 286–291. doi: 10.1016/j.yhbeh.2006.11.005
- Wedekind, C., and Furi, S. (1997). Body odour preference in men and women: do they aim for specific MHC combinations or simply heterozygosity. *Proc. Biol. Sci.* 264, 1471–1479. doi: 10.1098/rspb.1997.0204
- Wedekind, C., Seebach, T., Bettens, F., and Paepke, A. J. (1996). MHC-dependent mate preference in humans. *Proc. Biol. Sci.* 260, 245–249.
- Winternitz, J., Abbate, J. L., Huchard, E., Havlíček, J., and Garamszeg, Z. (2017). Patterns of MHC-dependent mate selection in humans and nonhumans primates a meta-analysis. *Mol. Ecol.* 26, 668–688. doi: 10.1111/mec.13920
- Zuniga, A., Stevenson, R. J., Mahmut, M. K., and Stephen, I. D. (2017). Diet quality and the attractiveness of male body odor. *Evol. Hum. Behav.* 38, 136–143. doi: 10.1016/j.appet.2015.11.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a past co-authorship with one of the authors RS.

Copyright © 2019 Mahmut and Stevenson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Interplay Between Economic Status and Attractiveness, and the Importance of Attire in Mate Choice Judgments

Amany Gouda-Vossos¹, Robert C. Brooks¹ and Barnaby J. W. Dixon^{2*}

¹ Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, NSW, Australia, ² School of Psychology, The University of Queensland, Brisbane, QLD, Australia

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Viktoria Mileva,
University of Stirling, United Kingdom
Andrew Thomas,
Swansea University, United Kingdom

*Correspondence:

Barnaby J. W. Dixon
b.dixon@uq.edu.au

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 05 September 2018

Accepted: 15 February 2019

Published: 21 March 2019

Citation:

Gouda-Vossos A, Brooks RC and
Dixon BJW (2019) The Interplay
Between Economic Status
and Attractiveness, and
the Importance of Attire in Mate
Choice Judgments.
Front. Psychol. 10:462.
doi: 10.3389/fpsyg.2019.00462

Desirable characteristics of “opposite sex others,” such as physical attractiveness and economic status, can influence how individuals are judged, and this is different for men and women. However, under various social contexts where cues of higher or lower economic status is suggested, sex differences in judgments related to mate choice have not been fully explored. In two studies, ratings of economic status and attractiveness were quantified for male and female targets that were presented under various social contexts. Study 1 assessed judgments ($n = 1,359$) of images of nine male and nine female targets in different sized groups containing only opposite-sex others (i.e., group size). While we found no significant effects of group size on male and female attractiveness, target female economic status increased when surrounded by two or more men. An *ad hoc* analysis controlling for the attire of the targets (business or casual) found that the association between target female economic status and group size occurred when females were in business attire. Study 2 investigates this effect further by presenting images of 12 males and 12 females, in higher and lower status attire (i.e., business and casual clothing) and measured judgments of attractiveness and economic status among women and men ($n = 1,038$). Consistent with the results of Study 1, female economic status was only affected when women were in business attire. However, female economic status decreased when in the presence of other men in business attire. There were no sex differences in judgments of economic status when judging stimuli in casual attire. Additionally, negative associations between attractiveness and economic status were found for males presented in casual attire. We discuss these results in the light of evolutionary sexual conflict theory by demonstrating how the asymmetrical importance of status between men and women can influence mate choice judgments.

Keywords: sex difference, status, attire, attractiveness, mate choice copying, economics

GENERAL INTRODUCTION

Evolved mate preferences often target attributes that signal dimensions of reproductive health (Buss and Schmitt, 1993; Puts, 2010). In women, age-related physical cues such as feminine facial shape, breast morphology, and an hourglass distribution of body fat are attractive to men (Jasińska et al., 2004; Singh et al., 2010; Dixon et al., 2011, 2015; Marcinkowska et al., 2014), ostensibly because they signal fecundability. In men, muscularity, vocal pitch, and facial masculinity provide

information regarding health, age, social status, dominance, and formidability, that enhance mating success (Archer, 2009; Puts, 2010; Hill et al., 2013; Dixon et al., 2014).

Judgments of physical attractiveness are also shaped by factors other than physical attributes (Dixon, 2019; Luoto, 2019). For instance, men are more likely to be rated as more physically attractive when they are presented with high status cues such as an expensive car (Dunn and Searle, 2010; Shuler and McCord, 2010) or an upscale apartment (Dunn and Hill, 2014). While these cues may not influence ratings of women's physical attractiveness, high status may drive intersexual competition between women (Wang and Griskevicius, 2013). Judgments of physical attractiveness also increase with the addition of other people, an effect known as "mate choice copying" (Waynforth, 2007). Men are more likely to be rated as more attractive and to have higher economic status when in the presence of women, whereas mate choice copying effects are negligible when women are in the presence of men (Gouda-Vossos et al., 2016, 2018).

The associations between sex, economic status, and physical attractiveness may reflect evolved sex differences in mate choice (Gouda-Vossos et al., 2018). Throughout human evolution, resource acquisition positively influenced male reproductive success (Low, 1990, 2000; Smith, 2004), so that sexual selection may have favored status seeking behavior in men (Betzig, 1986; Dixon, 2016; Von Rueden and Jaeggi, 2016). An association between men's status and reproductive success has been reported across many small-scale (Von Rueden and Jaeggi, 2016) and several industrialized societies (Li et al., 2002; Hopcroft, 2006), which may have implications regarding the formation of social perceptions of women and men. For instance, participants judged female economic status comparatively lower than that of the males they were presented alongside (Gouda-Vossos et al., 2016). Unlike men, women's success can be judged negatively, as high status or successful women are more frequently derogated (Heilman et al., 2004), particularly when dressed in short skirts or shirts displaying cleavage (Glick et al., 2005; Howlett et al., 2015). Conversely, physical attractiveness is beneficial to women as attractive individuals receive favorable treatment (Rosenblat, 2008) and are more likely to find jobs and get promoted (Hamermesh and Biddle, 1993; Pfann et al., 2000), which benefits women more than men within employment scenarios (Busetta et al., 2013). However, the interplay between physical attractiveness, economic status, and attire in mixed social contexts that are more comparable to real-world scenarios has yet to be explored.

The current research assesses how modifiable cues of economic status influence judgments of an individual's physical attractiveness and economic status. Previous studies reported that high status men were more likely to receive respect and praise compared to women in high status roles (Forsythe, 1990; Brase and Guy, 2004). Studies have also shown that high status women within mixed sex groups were just as likely as men to attain leadership positions (Goktepe and Schneier, 1989). However, high status women were judged to be less attractive and approachable than high status men (Howlett et al., 2013). Additionally, ratings of women's economic status are lower relative to men they are presented alongside

(Gouda-Vossos et al., 2016). However, whether this association persists when women are presented as higher in economic status than men or if physical attractiveness influences ratings of economic status remains unknown.

Based on evolutionary theories regarding the importance of status in male reproductive success (Hopcroft, 2006; Von Rueden et al., 2010) and mate choice copying theory (Waynforth, 2007; Gouda-Vossos et al., 2018), we predicted that cues of higher social status should have a stronger positive effect on ratings of economic status and physical attractiveness in men than in women. We also predicted that women's economic status would be rated lower than the men they were presented alongside, even if women appeared to be higher in economic status than men (Gouda-Vossos et al., 2016). We conducted two studies, both of which manipulated economic status via clothing in male and female stimuli and measured participant attractiveness and monetary earnings ratings of the stimuli. We first tested the effects of the presence of "opposite sex others" by manipulating mixed sex group sizes (Study 1). Based on the results of Study 1, we designed Study 2 wherein various forms of attire were used to manipulate social status, including the presence and absence of "opposite sex others" in various forms of attire.

STUDY 1: JUDGEMENTS OF ATTRACTIVENESS AND ECONOMIC STATUS WITHIN MIXED-SEX GROUPS

Dynamics within groups can vary depending on the size of social groups and the distributions of gender therein. All-male groups tend to be more aggressive and competitive toward other group members than all-female groups (Schopler et al., 2001). Additionally, all-male groups form more stable hierarchies faster than all-female groups (Anderson et al., 2001) and are more likely to collaborate intra-sexually when an outside threat is present (Vugt et al., 2007). Sharing cooperatively produced resources is also an important factor within collective groups (Mangel, 1990; Melis and Semmann, 2010) and the dynamics within groups can vary if some members are more likely to obtain larger portions of resources (relative to other members) and subsequently gain direct benefits (Williams, 2002). Among men, resource acquisition and holding potential enhance mating opportunities and mating success (Betzig, 1986; Von Rueden et al., 2010). As a result, expectations and opportunities vary between men and women within mixed-sex groups (Eagly and Johnson, 1990). Mate choice copying studies have also found that men are judged to be more physically attractive when presented within a group of women, while women presented alongside men are not (Milonoff et al., 2007; Hill and Buss, 2008; Dunn and Doria, 2010).

Behaviors within groups may be driven by similar mechanisms associated with sex differences in mate choice, especially in reference to associations between status (social or economic) and physical attractiveness. In men, cues of social status, dominance, and formidability that enhance male physical attractiveness (Hill and Buss, 2008; Archer, 2009; Puts, 2010; Dixon et al., 2017, 2019) and mating success (Hill et al., 2013;

Kordsmeyer et al., 2018) may also predict assertiveness and group leadership (Anderson et al., 2001; Geniole et al., 2015). While cues of status may also predict the emergence of female leaders in groups (Anderson et al., 2001) they may not augment women's physical attractiveness and mating success (Puts, 2010). To test the effects of social group size on ratings of male and female attractiveness and economic status, we presented images of women and men in the presence of social groups varying in the number of opposite sex targets. Thus, each male and female was rated alone and again alongside opposite sex targets in increments of 1, 2, and 4 additional opposite sex individuals.

Materials and Methods

Participants

Participants were recruited via Facebook, Twitter, and internal student email lists within the research institution, resulting in 1359 participants in total. All participants were over 18 years old and were not aware of the purpose of the study. Each participant provided details of their biological sex, age, and sexual orientation using a Kinsey Scale (Kinsey et al., 1948, 1953). As sexual orientation impacts on judgments of attractiveness of opposite sex targets (Pettersson et al., 2015, 2016, 2018; Valentova et al., 2017), only participants who were heterosexual or bisexual were retained in the analyses (i.e., Kinsey scale 0–3). The final analysis included 569 (women = 494; men = 75) participants who completed surveys including target males and 598 (women = 357; men = 241) who completed the survey's including target females. The age range of participants was $29.04 \text{ years} \pm 9.4$. The majority of participants listed their country of origin as Australia (38.7%), followed by USA (20%), then UK (17.7%). The majority identified as North Western European, British, or Irish (51.2%), followed by European Mixed Race (13.8%), then Southern European (3.8%) with 8% stating they "did not wish to report ethnicity."

Stimuli

Images of nine male and nine female targets surrounded by four members of the opposite sex (females and males, respectively) were chosen from a stock photo website¹. This resulted in a total of 18 original images. Each target was presented in four group size conditions [alone, one opposite sex, two opposite sex, and four opposite sex others; for examples, see **Electronic Supplementary Material S1** (ESM 1)]. Overall, 72 images were constructed and used in this study, with the target pose and facial expression identical between treatments. The targets ages ranged from 22 to 56 years (males: mean = 40, $SD \pm 12.3$, females: mean = 38, $SD \pm 12.7$). All images were professionally taken under standardized lighting and filters. Photographs were taken in workplaces and casual settings with positions of targets and opposite sex others varying from image to image.

Procedure

Experiments were conducted on-line via www.socialsci.com. Each participant entering the study was randomly assigned to one of four experiments in which they rated either male or female targets for either attractiveness or monetary earning

(i.e., economic status). The number of participants for each experiment was as follows: Attractiveness/target female: 202 women, 150 men; Earnings/target female: 155 women, 91 men; Attractiveness/target males: 259 women, 39 men; Earnings/target males: 235 women, 36 men. The study employed a "Within Target – Between Treatment" design where participants saw all nine targets in random order with the treatment (target alone, one opposite sex other, two opposite sex others, and four opposite sex others) for each target drawn at random. Similar designs have been used in past research on physical attractiveness (Janif et al., 2014, 2015; Brooks et al., 2015; Dixson et al., 2016). This research was approved by the University of New South Wales Human Research Ethics Advisory Board (Psychology) (HREAP 1880).

Participants were informed that they would be shown a range of images of people. In each image, the target was indicated with an arrow. If assigned to rate physical attractiveness, participants were asked to rate each target using a percentile scale from 0 to 100 where "50" indicated that the individual is more physically attractive than 50% of other individuals of the same sex (i.e., of median attractiveness). If rating economic status, participants were asked to rate each target using a percentile scale from 0 to 100 where "50" indicated the individual earns more than 50% of other same sex individuals in full time work (i.e., median income in full-time work).

Analysis

Multilevel modeling was used where data were organized so that each row represented one participants rating of one target in one treatment. Using the statistical software SPSS, separate general linear mixed models (MLMs) were fitted for the two dependant variables (physical attractiveness or economic status). In each of these models, Model ID was a repeated-measures factor, Participant ID was a random factor. Participant sex and Group Size (alone, +1 opposite sex individual, +2 opposite sex individual, and +4 opposite sex individuals) were included as fixed factors. SPSS does not calculate effect sizes for mixed models. Thus, we calculated approximate effect sizes as partial Eta-squared, from the *F*-test and degrees of freedom, although this practice has not been formally validated for multi-level models. When interpreting effect sizes, by convention, effects of 0.2, 0.5, and 0.8 are interpreted as small, medium, and large effect sizes, respectively.

Results

The Effect of Group Size on Male and Female Attractiveness and Economic Status

There were no significant main effect or interactions involving Group Size; suggesting no differences in the ratings of attractiveness across treatment (**Figures 1A,B** and **Table 1a**). The significant main effects of participant sex on ratings of target females were due to women rating target females as more attractive (mean = 59.85, $SE \pm 0.391$) than men (mean = 58.66, $SE \pm 0.456$).

Like the results for attractiveness, male ratings of economic status were not affected by Group Size as there was no significant main effect or interactions with participant sex (**Figure 1C** and **Table 1b**). However, ratings for target females revealed

¹ www.peopleimages.com

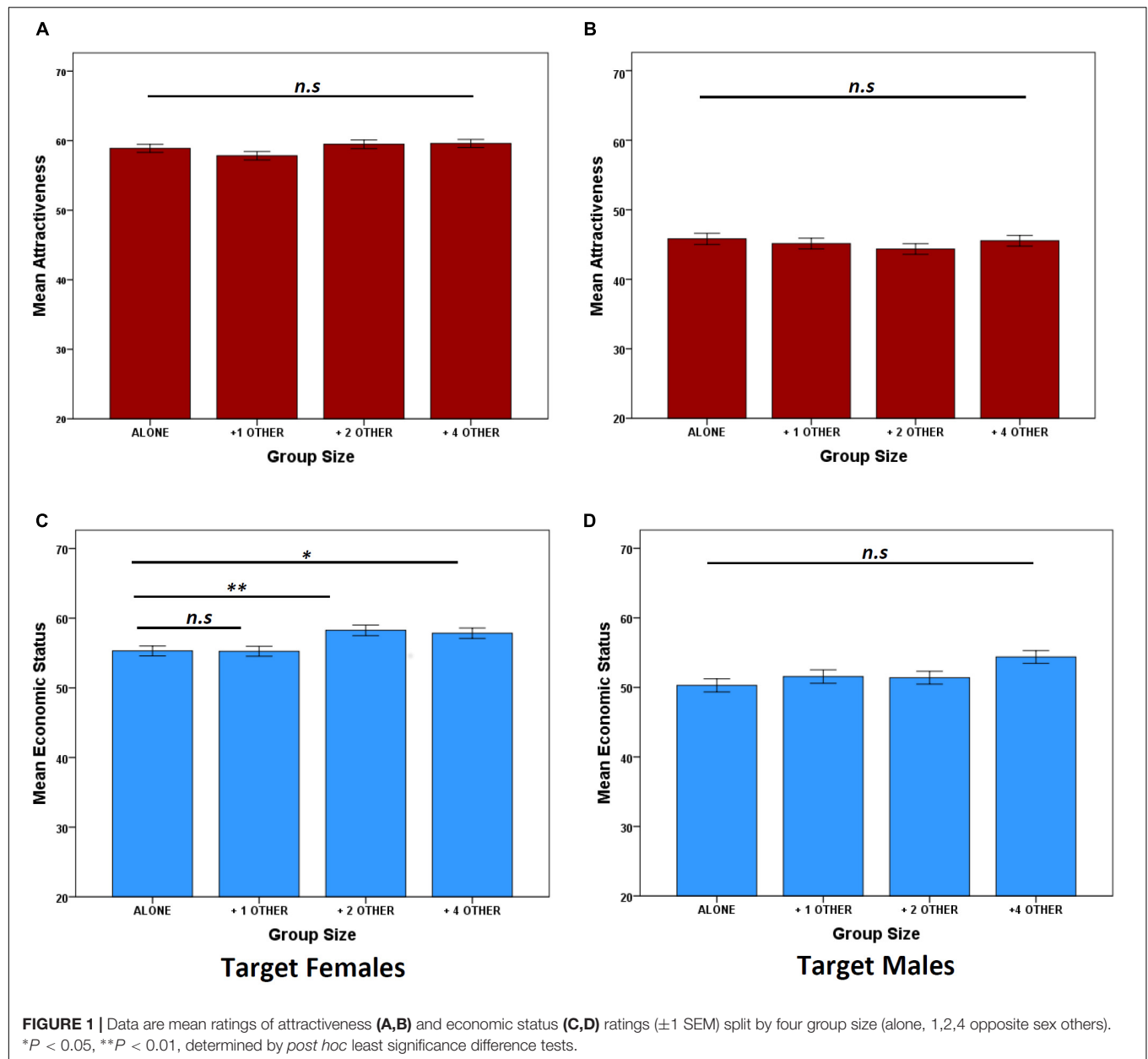


TABLE 1 | MLMs for target male and female rated attractiveness and economic status.

	Target female			Target males				
	<i>df</i>	<i>F</i>	<i>P-value</i>	η_p^2	<i>df</i>	<i>F</i>	<i>P-value</i>	η_p^2
(a) Attractiveness								
Akaike Information Criterion	26,858.21				23,691.96			
Group size	3, 3,074.69	1.74	0.156	0.002	3, 2,569.39	0.12	0.951	<0.001
Participant sex	1, 3,085.56	3.90	0.048	0.002	1, 2,597.86	1.11	0.291	<0.001
Group size * participant sex	3, 3,074.69	0.21	0.893	<0.001	3, 2,569.39	0.97	0.406	<0.001
(b) Economic status								
Akaike Information Criterion	18,833.71				22,008.02			
Group size	3, 2,084.79	4.50	0.004	0.006	3, 1,696.49	1.51	0.210	0.003
Participant sex	1, 2,109.86	1.66	0.198	<0.001	1, 1,730.98	3.06	0.080	0.002
Group size * participant sex	3, 2,084.79	0.35	0.788	<0.001	3, 1,696.49	0.23	0.874	<0.001

a significant main effect of Group Size (**Table 1b**) as female economic status increased incrementally with the addition of two and four males (**Figure 1D**). There was no significant Participant Sex \times Group Size interaction, suggesting that men and women were rating target females similarly (**Table 1b**).

The Effect of Attire and Group Size on Target Attractiveness and Economic Status; an *ad hoc* Analysis

Images included targets in either casual or business attire, which may have affected ratings. To test this, targets were classified as wearing either business or casual attire using methods from Forsythe (1990). Business attire referred to dark, angular, traditional business suits whereas casual attire referred to light, informal, everyday wear. There were four business and five casual attired target males and five business and four casual attired target females. Full analysis can be found in **Electronic Supplementary Materials S2, S3** (ESM 2: Male and Female Attractiveness and ESM 3: Male and Female Economic Status).

Another series of MLMs were conducted, with attire included as a fixed factor. We found no effects of Attire on target male attractiveness ($F_{1,2555} = 0.851$, $P = 0.356$, $\eta_p^2 = 0.00033$), Group Size ($F_{3,2555} = 0.091$, $P = 0.965$, $\eta_p^2 = <0.001$), and Participant Sex ($F_{1,2555} = 0.961$, $P = 0.339$, $\eta_p^2 = <0.001$). The effect sizes for all main effects were small for male attractiveness (i.e., less than 0.2), suggesting that both Attire and Group Size do not strongly impact on male attractiveness.

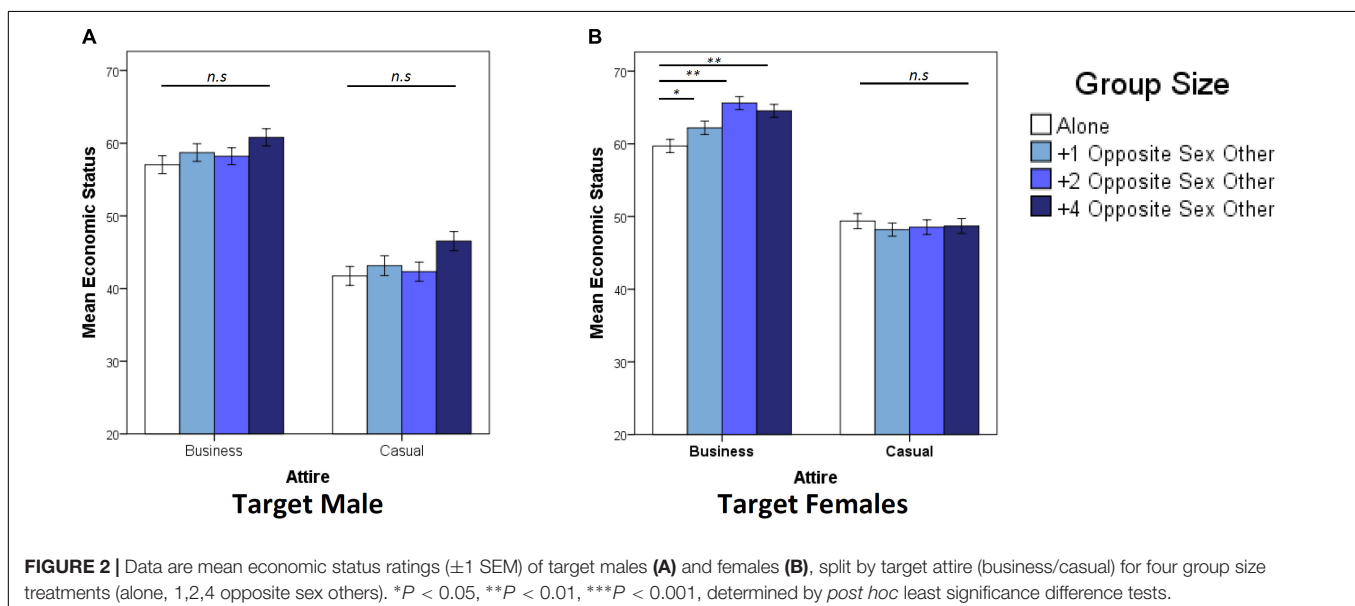
There was a main effect of Attire on target male economic status ($F_{1,1877} = 533.83$, $P < 0.001$, $\eta_p^2 = 0.220$), but not Group Size ($F_{3,1853} = 1.408$, $P = 0.239$, $\eta_p^2 = 0.002$), or Participant Sex ($F_{1,1877} = 2.156$, $P = 0.142$, $\eta_p^2 = 0.001$). Both men and women rated male economic status higher when in business attire (mean = 65.32, $SE \pm 0.713$) than when in casual attire (mean = 39.43, $SE \pm 0.865$). We did not find mate choice copying effects, which suggest that the type of attire men were

wearing influences ratings of target males more than the presence of other females.

There was a main effect of Attire on attractiveness ratings of target females ($F_{1,2963} = 68.13$, $P < 0.001$, $\eta_p^2 = 0.023$) but no main effect of Group Size ($F_{3,2952} = 1.937$, $P = 0.121$, $\eta_p^2 = 0.002$) or Participant Sex ($F_{1,2963} = 2.982$, $P = 0.084$, $\eta_p^2 = 0.010$). Target females were rated as more attractive when in business (mean = 61.46, $SE \pm 0.404$) than Casual Attire (mean = 56.52, $SE \pm 0.442$), although the effect size was small (i.e., below 0.2) and comparable to Group Size, suggesting that the impact of Attire on female attractiveness is small.

Ratings of female earnings were also significantly affected by Attire ($F_{1,2075} = 356.54$, $P < 0.001$, $\eta_p^2 = 0.150$) and Group Size ($F_{2,2057} = 3.181$, $P = 0.023$, $\eta_p^2 = 0.005$), although both effect sizes were small (i.e., below 0.2). A significant Attire \times Participant Sex interaction ($F_{1,2057} = 18.68$, $P < 0.001$, $\eta_p^2 = 0.009$) occurred due to women rating females higher in economic status when in business attire than men (women mean = 63.57, $SE \pm 0.566$; men mean = 61.6, $SE \pm 0.740$; $P = 0.043$) and lower when in casual attire than men (women mean = 47.58, $SE \pm 0.622$; men mean = 51.89, $SE \pm 0.785$; $P < 0.001$). There was also a significant Attire \times Group Size interaction ($F_{1,2057} = 6.369$, $P < 0.001$, $\eta_p^2 = 0.009$), which reflects that ratings of female economic status increased incrementally with the addition of two and four males when females were presented in business but not casual attire (**Figures 2A,B**). This suggests that the original results of female economic status were likely driven by the responses toward target females in business attire, as opposed to casual attire.

In order to determine the model of best fit, we used the Akaike Information Criterion (AIC). We monitored the AIC between the original models and the *ad hoc* analysis. If AIC changes downward by more than 2 units, the model is significantly better (Bozdogan, 1987). Using both the level of significance of interactions and the AIC allows us to test the validity of different restrictions of a model and to choose a



model with the smallest probability of rejection to be the best fitting model as opposed to choosing based on a *priori* ground (Bozdogan, 1987). Across all four *ad hoc* models, the AIC was significantly lower than the original models constructed (Female Attractiveness, Original AIC: 26858.21, *ad hoc* AIC: 26,768.36; Male Attractiveness, Original AIC: 23,691.96, *ad hoc* AIC: 23,657.82; Female Earnings, Original AIC: 18,833.71, *ad hoc* AIC: 18,405.63; Male Earnings, Original AIC: 22,008.02, *ad hoc* AIC: 21,434.76). This suggests that entering Attire as a fixed factor improves all the models.

Discussion

Contrary to our predictions, ratings of attractiveness for target male and females were not influenced by social group size. Further, the economic status of males was unaffected when in the presence of an opposite sex other (i.e., female). Whereas the addition of opposite sex others positively influenced the economic status of target females. Previous studies have found that woman's task proficiency (Balkwell and Berger, 1996), economic status (Gouda-Vossos et al., 2016), and social status (Eagly and Karau, 2002) were rated lower than the men they were compared with, suggesting that the mere presence of a man can lower perceptions of women's status within an economic hierarchy. However, the type of attire women wear may reduce these effects as women dressed in more masculine attire (i.e., traditional business attire) are seen as having better managerial characteristics (Forsythe et al., 1984, 1985) are more likely to get hired for leadership positions (Forsythe, 1990), and are just as likely as men to emerge as leaders within a mixed sex group (Goktepe and Schneier, 1989). Taken together with the results of the current studies, women's perceived economic status appears to be heavily influenced by high status and masculine cues such as business attire and the number of men within their immediate presence.

However, we are limited in making these assumptions regarding attire, as the numbers of targets used were too small after separation for *ad hoc* analysis (i.e., four business and five casual attired target males and five business and four casual attired target females). Further, the position of the target in each photo was not randomized or controlled and therefore we could not conclude whether participants perceived targets as leaders, which could have influenced ratings of attractiveness and economic status. It is also possible that even though female economic status increased with the addition of "male others," ratings of female status may still be made relative to men. Unfortunately, we did not obtain ratings of earnings and attractiveness of the male "opposite sex others" to confirm this. Thus, we designed a second study focused on the impact of attire of various social status (business/casual) and measured effects of male and female attractiveness and economic status in targets presented individually and when paired (Study 2).

STUDY 2: THE EFFECTS OF ATTIRE ON MEN AND WOMEN

Attire communicates information relating to identity, social status, and position within a hierarchy (Roach-Higgins and

Eicher, 1992). Molloy and Potter (1975) suggested that attire may play a pivotal role in judgments of an individual's credibility, likeability, interpersonal attractiveness, and dominance. Bassett et al. (1979) found that high status clothing positively influenced judgments of credibility. However, females were rated lower than males across all four measurements that composed credibility (i.e., potency, character, composure, and competence). Unlike men, women's success can be judged negatively, as high status or successful women are more frequently derogated (Heilman et al., 2004) and are judged more negatively when dressed in short skirts (Glick et al., 2005; Howlett et al., 2015). However, physically attractive women receive better treatment (Rosenblat, 2008), are more likely to find employment, and are more likely to get promoted (Hamermesh and Biddle, 1993; Pfann et al., 2000), which may not be the case among men (Busetta et al., 2013).

In Study 2, we measured associations between rated physical attractiveness and economic status in male and female targets in different attires. We also assessed if the presence of opposite sex others in various attire influenced judgments of male and female targets. Based on mate choice copying theories (Waynforth, 2007; Gouda-Vossos et al., 2018), we hypothesized that males would attain high ratings of physical attractiveness and economic status when presented in high status attire (i.e., business attire) regardless of the attire the opposite sex other. As economic status may not have had strong effects on female reproductive success during human evolution (Betzig, 1986), we did not predict a positive association between attractiveness and economic status with target females. Based on the results of Study 1, we predicted that female economic status will be limited to the ratings of males when in casual but not business attire.

Materials and Methods

Participants

A total of 1,035 participants were recruited. All participants were over 18 years old and were not aware of the purpose of the study. Each participant provided details of their biological sex, age, and sexual orientation using the Kinsey Scale (Kinsey et al., 1948, 1953). Participants received \$1US. As in Study 1, only participants who were heterosexual or bisexual were retained in the analyses (i.e., Kinsey scale 0–3). A total of 459 females and 578 males (Age 32 ± 10.5) were included in the final analysis. The majority of participants listed their country of origin as United States (82.4%), followed by Southern Europe (4.4%), then Australia (3.4%). The majority identified as ethnically North Western European, British, or Irish (40.4%), followed by European Mixed Race (17.2%), Southern European (7.2%), and 10% elected not to state their ethnicity.

Stimuli

Full body, color photographs of 12 male and 12 female targets were obtained from a stock photo website². All photographs were taken using standardized lighting and filters and were on a white background. Two sets of images for each male and female target were obtained (i.e., either in casual attire or business attire),

²www.peopleimages.com

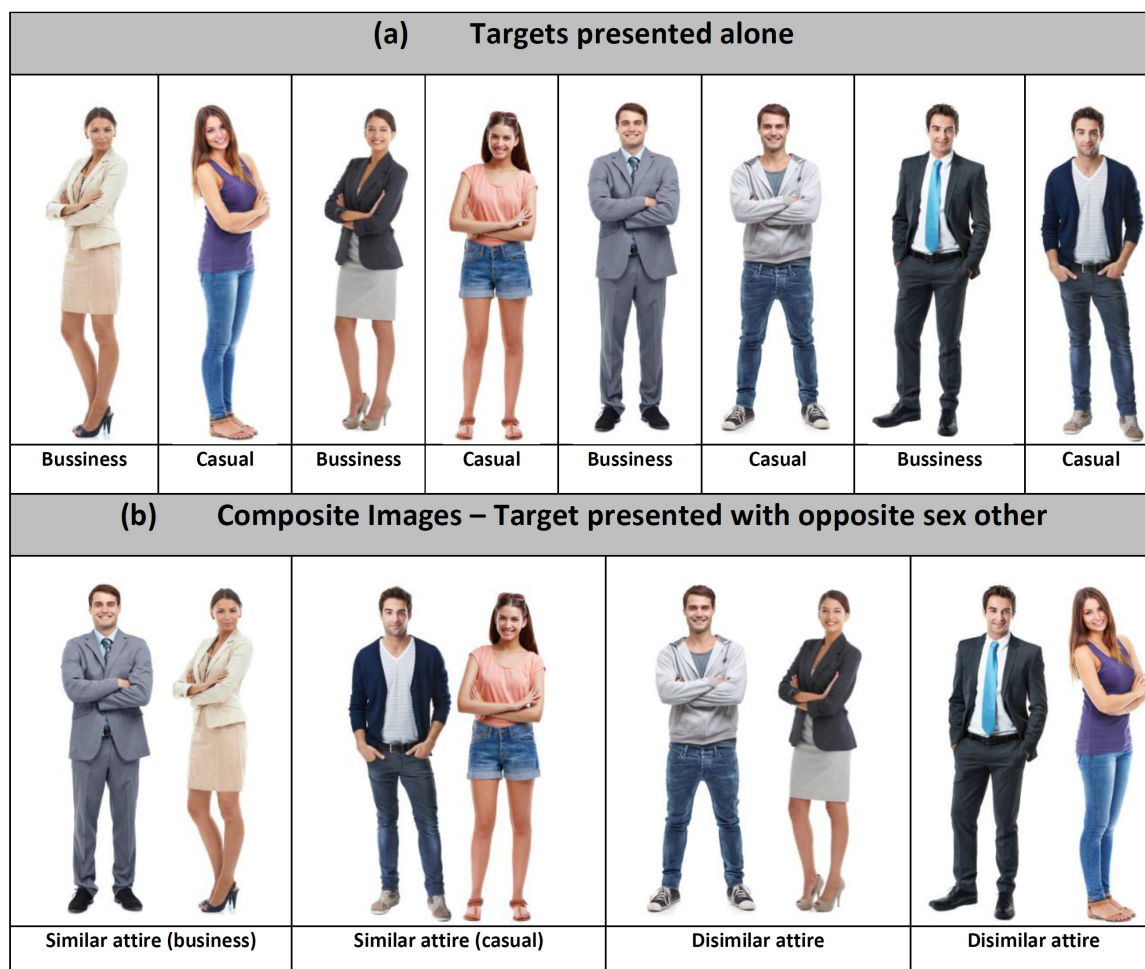


FIGURE 3 | Example images used for targets presented alone **(A)** and with opposite sex others **(B)**. Film Media courtesy of Yuri Arcurs Photography Aps, Used by Permission.

comprised of 24 male and 24 female target images. We then created composite images, where each target (in both business and casual attire) was paired with opposite sex others (six in business and six in casual attire), resulting in a total of 144 composite images for male targets and 144 composite images for female targets (**Figure 3**). Business attire included suits, collared shirts, and pencil skirts (for females not wearing pant suits). Casual attire included t-shirts, jeans, shorts, or skirts (**Figure 3**). PeopleImages.com collects information on the targets they recruit, including biological sex, age, and ethnicity. The targets ages ranged from 20 to 30 years (Males Target mean = 25, $SD \pm 4.8$ years; Females Target Mean = 24, $SD \pm 4.3$) and the majority were Caucasian (66%) followed by multi-ethnic (16%), then African and Latino (9% each).

Procedure

Studies were conducted online using the SocSci platform³ and participants were recruited via MTurk. Participants each

rated two batches of 12 images (24 in total). One batch of 12 images included either male or female targets alone. The other batch included male or female targets with opposite sex others. The order of the batches as well as the target sex was fully randomized, so that participants were presented with four possible combinations (i.e., female alone/male with other; female with other/male alone; male alone/female with other; and male with other/female alone). Within each batch, each target was drawn and shown once, in random order and either in casual or business attire (drawn at random with equal probability) (see **Figures 3A,B** for examples). Thus, each target was presented to the participant either alone or with an opposite sex other, in either business or casual attire. This design ensured participants saw all possible targets (male and female) in only one type of attire (i.e., Within Target – Between Treatment design). By designing the study in this manner, participants do not see the same target more than once in different scenarios, which minimizes possible carry-over effects and avoided participants deciphering the true nature of the study that previous studies have shown to influence ratings of targets (Chen, 2008). Experimental designs like this

³ www.socialsci.com

have been previously employed to test preferences for physical appearance (Janif et al., 2014, 2015; Brooks et al., 2015; Dixson et al., 2016). Participants rated targets using a sliding scale (from 0 to 100) provided below each image for physical attractiveness and economic status using the same scales as in Study 1. This research was approved by the University of New South Wales Human Research Ethics Advisory Board (HREA 155047).

Analysis

Using the statistical software, SPSS, we first tested the influence of individual sex and attire on ratings of attractiveness and economic status by focusing on the ratings of attractiveness and economic status of targets when presented alone. This allowed us to determine how target attire (business/casual) and target sex (female/male) influence these ratings. This 2×2 between-subject design (Sex of target – Male/Female) \times (Attire of Target – business/casual) employed MANOVAs, with rated attractiveness and economic status as dependant variables.

We then assessed if the attire of the opposite sex other influenced male and female attractiveness and economic status by focusing on ratings of targets when presented with opposite sex others. This was a $2 \times 2 \times 2 \times 2$ between-subject design (Sex of target – Male/Female) \times (Attire of TARGET – business/casual) \times (Attire of OTHER – business/casual) \times (Participant Sex – Men/Women). We analyzed the results of targets when presented with an opposite sex other of varying attire (business/casual), using separate general linear mixed models (MLM) for ratings of physical attractiveness and economic status for each study. Target ID, Subject ID, and OtherID were included as random factors to specify the covariance structure for the residuals. Target Attire, Target Sex, Other Attire, and Participant Sex were fixed factors. All main effects and interactions were assessed.

Results

Effects of Individual Attire and Sex on Attractiveness and Economic Status Ratings

The multivariate analysis revealed significant main effect of Target Sex, Target Attire, and their interaction (Table 2). For female targets, there were positive associations between Attractiveness and Economic status ratings when presented in both casual and business attire (Figure 4). In contrast, male targets were rated negatively for attractiveness and economic status when in casual attire (Figure 4). There was a significant

Target Sex \times Target Attire interaction (Table 2), which reflects target females were rated lower for economic status (target female mean = 54.03, $SD \pm 11.27$; target male = 55.33, $SD \pm 12.04$) but higher for attractiveness (target female = 67.178, $SD \pm 5.33$; target male = 56.32, $SD \pm 5.27$) than male targets. There was also a significant Target Sex \times Target Attire interaction (Table 2), so that target females were rated as more attractive in casual attire (mean = 69.28, $SD \pm 4.72$) than business attire (mean = 65.08, $SD \pm 5.91$).

The Influence of Target Attire and the Attire of Opposite Sex Other on Target Attractiveness and Economic Status

There was a significant main effect of Target Sex (Table 3a), so that female targets were rated as more attractive than male targets (Target Female mean = 67.32, $SE \pm 1.20$; Target Male Mean = 56.32, $SE \pm 1.16$; $P < 0.001$). The Target Sex \times Participant Sex interaction was not statistically significant (Table 3a). There was a significant Target Attire \times Target Sex \times Participant Sex interaction for attractiveness ratings that was driven by the ratings from men, who rated female attractiveness lower when in business attire than casual; and male attractiveness higher in business than casual (Table 3a and Figure 5). There were no significant effects due to women's ratings (Figure 5).

There were no significant main effects or interactions involving Other Attire on rated attractiveness (Table 3a and Figure 6A). However, there was a significant Other Attire \times Target Attire interaction for ratings of economic status (Table 3b and Figure 6B). Targets in business attire presented alongside “others” in business attire were rated higher in economic status than when presented alongside “others” in casual attire (Figure 6B), which did not vary with Target Sex (Table 3b). A significant Target Sex \times Target Attire interaction reflected men in business attire were rated significantly higher for economic status than women in business attire (Table 3b and Figure 7). Ratings of males and females in casual attire did not differ significantly (Figure 7). There were no significant main effects or interactions involving participant sex (Table 3b). For additional analyses see Supplementary Tables S4, S5.

Discussion

As predicted, economic status ratings were higher when male and female targets were presented in business than casual attire.

TABLE 2 | MANOVA for rated economic status and attractiveness for targets presented alone.

Pillai's trace	Within subject effects											
	Target sex				Target attire				Target sex * target attire			
	<i>F</i>	<i>df_n</i>	<i>P</i>	η^2_p	<i>F</i>	<i>df_n</i>	<i>P</i>	η^2_p	<i>F</i>	<i>df_n</i>	<i>P</i>	η^2_p
MANOVA	202.22	2,283	<0.001	0.588	895.93	2,283	<0.001	0.864	9.39	2,283	<0.001	0.062
Between subject effects												
Economic status	5.358	1,284	0.021	0.019	1426.96	1,284	<0.001	0.834	2.15	1,284	0.144	0.008
Attractiveness	301.98	1,284	<0.001	0.515	5.706	1,284	0.018	0.020	18.76	1,284	<0.001	0.620

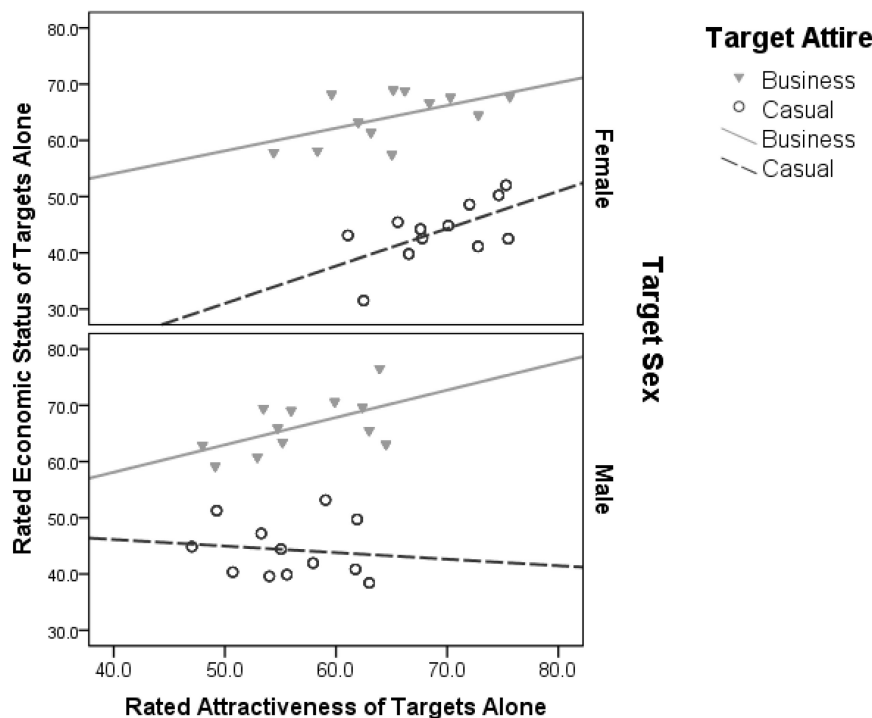


FIGURE 4 | The association between rated attractiveness and economic status of target males and females when presented alone in business and casual attire.

TABLE 3 | MLMs for target male and female rated attractiveness and economic status.

	(a) Rated attractiveness				(b) Rated economic status			
AIC without factors (random factors only)	54,248.21				51,332.19			
AIC including factors	54,073.10				49,135.69			
	<i>F</i>	<i>df</i>	<i>P</i>	η^2_p	<i>F</i>	<i>df</i>	<i>P</i>	η^2_p
Target attire	1.69	1, 6056.64	0.193	<0.001	2600.64	1, 5520.90	<0.001	0.320
Target sex	87.27	1, 540.25	<0.001	0.139	9.68	1, 489.30	0.002	0.019
Other attire	1.39	1, 6050.47	0.238	<0.001	1.72	1, 5561.22	0.189	<0.001
Participant sex	4.30	1, 540.23	0.039	0.008	0.82	1, 489.30	0.365	0.002
Target attire * target sex	34.15	1, 6054.93	<0.001	0.006	47.32	1, 5520.88	<0.001	0.009
Target attire * others attire	0.69	1, 6045.44	0.405	<0.001	7.39	1, 5529.43	0.007	0.001
Target attire * participant sex	1.11	1, 6043.64	0.292	<0.001	2.050	1, 5521.28	0.152	<0.001
Target sex * others attire	1.52	1, 6045.69	0.218	<0.001	0.11	1, 5561.10	0.741	<0.001
Target sex * participant sex	0.74	1, 540.26	0.391	0.001	0.05	1, 489.31	0.815	<0.001
Others attire * participant sex	0.08	1, 6039.05	0.778	<0.001	0.32	1, 5560.11	0.571	<0.001
Target attire * target sex * others attire	2.65	1, 6060.65	0.104	<0.001	0.03	1, 5528.28	0.863	<0.001
Target attire * target sex * participant sex	7.24	1, 6039.18	0.007	0.001	1.67	1, 5521.30	0.196	<0.001
Target attire * others attire * participant sex	0.88	1, 6055.73	0.349	<0.001	0.001	1, 5528.55	0.972	<0.001
Target sex * others attire * participant sex	1.90	1, 6040.29	0.168	<0.001	1.47	1, 5559.12	0.225	<0.001
Target attire * target sex * others attire * participant sex	1.53	1, 6055.86	0.216	<0.001	0.98	1, 5528.10	0.323	<0.001

AIC, Akaike Information Criterion.

Additionally, male and female economic status ratings increased when presented alongside others in business attire. In contrast to Study 1, target females received lower attractiveness ratings in business than casual attire, which was driven by men's ratings. The reduction in female attractiveness ratings when presented

in business attire is consistent with previous research reporting high status women were judged negatively, less attractive, and less approachable than lower status women (Bassett et al., 1979; Forsythe, 1990; Heilman et al., 2004; Lavin A. et al., 2009). We also report female economic status was rated lower than the

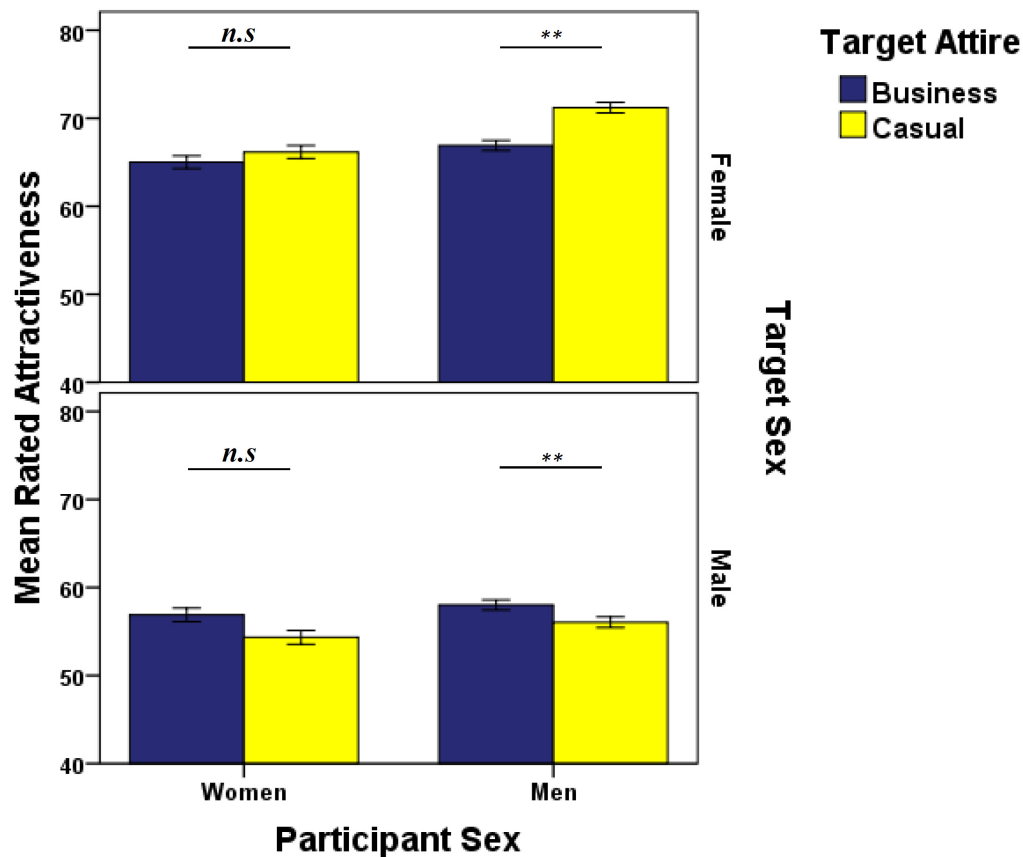


FIGURE 5 | Data are mean attractiveness ratings (± 1 SEM) from women and men (participant sex), split by target sex for two attire treatments (business/casual). **** $P < 0.01$, determined by *post hoc* least significance difference tests.

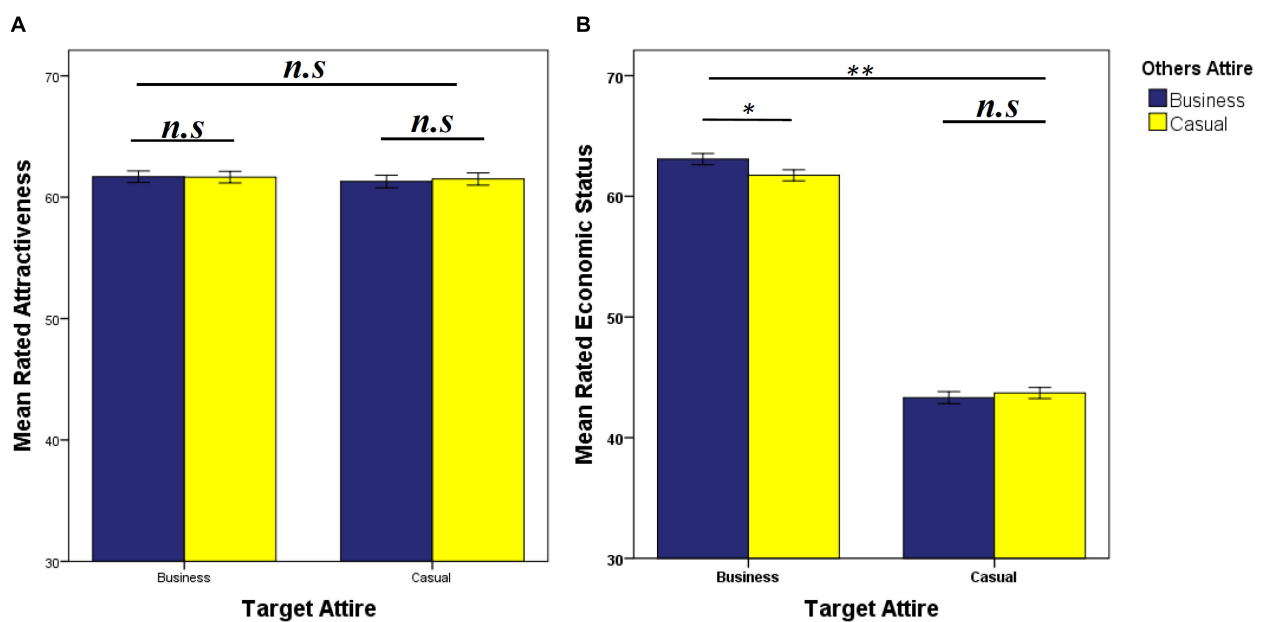
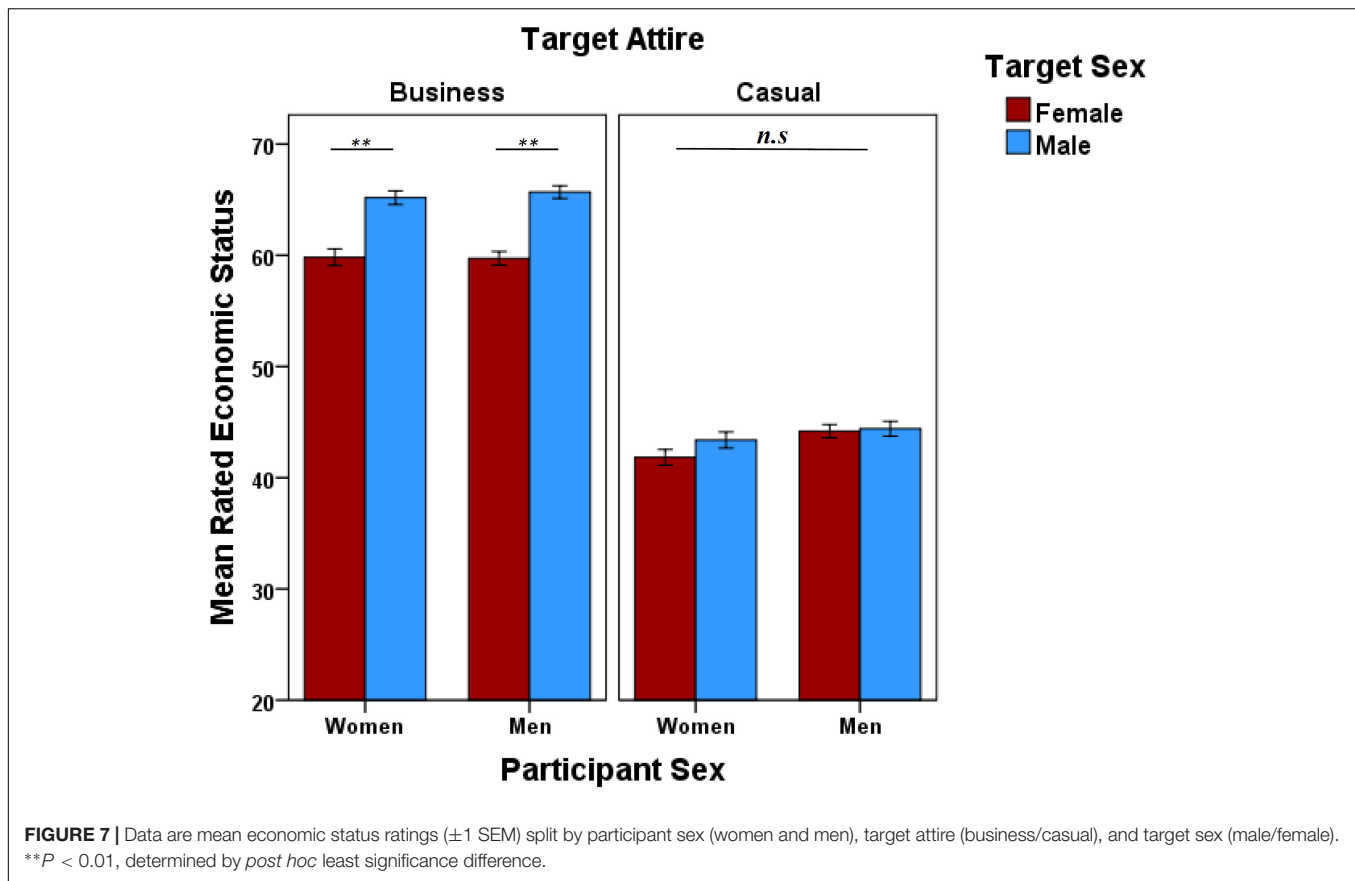


FIGURE 6 | Data are mean ratings (± 1 SEM) for (A) attractiveness and (B) economic status, split by other attire (business/casual) and target attire (business/casual). **P < 0.05*, ***P < 0.01*, determined by *post hoc* least significance difference tests.



men they were presented alongside, indicating that perceptions of women's economic status are influenced by the men they are presented with (Gouda-Vossos et al., 2016). Interestingly, the effects of economic status and sex disappeared when target males and females were presented in casual attire, so that sex differences in perceptions of status were specific to status-related social cues.

Our predictions regarding positive associations between attractiveness and economic status in male targets were supported, but only in males presented in business and not casual attire. Unexpectedly, we found positive associations between attractiveness and economic status ratings for women in both types of attire whereas previous research reported strong associations between economic status and attractiveness in males, with mixed results in women (Townsend and Levy, 1990; Hanson et al., 1991; Shuler and McCord, 2010). Our findings suggest that the influence of status-related clothing on judgments of attractiveness among women and men may be less robust than previously reported.

GENERAL DISCUSSION

A "Wall Street" article discussing attire and business practices stated that "traditional business dress is seen as a uniform... it

simplifies decision making and makes hierarchies easy to read." (Binkley, 2008). Our findings reinforce this sentiment, as participants made clear distinctions in physical attractiveness and economic status judgments based on clothing. We also report that women's economic status is judged relative to and lower than the men they are presented alongside. These sex differences in judgments of economic status disappeared when target males and females were presented in casual attire, demonstrating that judgments of women and men's economic status are most influenced by traditionally masculine clothing.

Status seeking is positively associated with men's mating and reproductive success (Betzig, 1986; Hopcroft, 2006) with high economic status associated more with ideals surrounding maleness and masculinity than femininity (Akerlof and Kranton, 2000). Past studies have found that, regardless of sex, group members expressing masculine gender roles or dress in masculine attire are more likely to emerge as leaders and are judged as more forceful and aggressive than those expressing outwardly feminine characteristics (Goktepe and Schneier, 1989; Forsythe, 1990). A limitation of the current study was that we did not compare the effects within same sex groups. In our previous study (Gouda-Vossos et al., 2016), we reported that the economic status of target males was judged to be highest when presented with another man. High status men form same-sex alliances and partnerships

(Von Rueden et al., 2010; Von Rueden, 2011), and the presence of two men not obviously in conflict may give the appearance that targets were forming same-sex partnerships. This was not the case when target females were presented alongside another female, as economic status was rated lower than when female targets were presented alone (Gouda-Vossos et al., 2016). In order to fully understand how men and women's economic status are perceived within various group dynamics, and if economic status is truly judged within a masculine hierarchy, comparisons within same-sex groups would be a worthwhile extension of the current research.

In male dominated social environments characterized by defined hierarchies (Anderson et al., 2001; Schmid Mast, 2004) business attire is associated with more masculine and socially dominant attributes (Forsythe, 1990). Without clothing that clearly communicates economic status, it may be difficult for people to judge where others fall within a hierarchy, which may be why economic status ratings were more neutral when targets were presented in casual attire. It was also unsurprising that female attractiveness was rated lower when presented in business attire. However, this directly contradicts studies reporting no negative influence on female physical attractiveness when presented with high status cues such as cars (Brase and Richmond, 2004) or luxury apartments (Dunn and Hill, 2014). This suggests that judgments of female attractiveness are more likely to vary when women are presented as being of higher status rather than alongside high status cues. It could be argued that participants did not believe that the women in the study actually own the high-status cues (i.e., cars, luxury apartments, etc.); with attire being a more convincing indicator of earned status. By presenting target females as high status individuals, it may communicate economic independence and decrease the attractiveness of female targets to men.

The current study also found positive associations between attractiveness and economic status among male and female targets, except for males presented in casual attire. Judgments of men's economic status and physical attractiveness are strongly positively correlated (Townsend and Levy, 1990; Hanson et al., 1991; Shuler and McCord, 2010) with competence, financial worth, and credibility being more consistently associated with men in business than casual attire (Bassett et al., 1979; Morris et al., 1996; Lavin A. M. et al., 2009). However, even subtle differences in attire within male-dominated business environments can lead to negative criticisms and attitudes toward men. For instance, men are perceived to be less confident, successful, and having lower salaries when presented in "off the peg" suits as opposed to "tailored suits" (Howlett et al., 2013). Further, men experience greater verbal harassment when presented in non-traditional business attire (i.e., business casual) than when presented in business attire (Kwantes et al., 2011). It could be argued that the economic status of more attractive male targets in casual attire was penalized in the current study, demonstrating how culturally malleable cues of status interplays with male attractiveness, possibly influencing women's mate preferences.

CONCLUSION

Maestripieri et al. (2017) marshaled a comprehensive review on financial and prosocial biases and concluded that attractive individuals, especially women, were more likely to attain financial benefits and better treatment than their less attractive counterparts. The results of the current studies demonstrated that high status individuals, especially men, receive more favorable judgments relating to mate choice (i.e., attractiveness). However, whether this leads to better treatment or increased financial gain remains to be fully explored. Ostensibly, men and women both benefit from being highly attractive or high status, however, this benefit is not distributed equally. Although this is consistent with ideals surrounding the asymmetrical importance of status in males and physical attractiveness in women within mating contexts, the results of the current studies reflect how this may lead to unfair judgments and, possibly, unfair treatment of both men and women.

AUTHOR CONTRIBUTIONS

AG-V, RB, and BD contributed conception and design of the study. AG-V carried out studies and organized the database. AG-V, RB, and BD performed the statistical analysis. AG-V wrote the first draft of the manuscript. AG-V, RB, and BD wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by the Australian Research Council awarded to RB and BD and a University of Queensland Post-Doctoral Fellowship to BD.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00462/full#supplementary-material>

MATERIAL S1 | Example images used for targets presented alone and with opposite sex others.

MATERIAL S2 | *Ad Hoc* MLMs of rated attractiveness of male and female models by men and women, factoring in 'target attire'.

MATERIAL S3 | *Ad Hoc* MLMs of rated economic status of male and female models by men and women, factoring in 'target attire'.

MATERIAL S4 | Data for groups study 1 separated by target sex and measure (attractiveness, earnings).

MATERIAL S5 | Data for attire study 2 separated by measure (attractiveness, earnings).

REFERENCES

- Akerlof, G. A., and Kranton, R. E. (2000). Economics and identity. *Q. J. Econ.* 115, 715–753. doi: 10.1162/003355300554881
- Anderson, C., John, O. P., Keltner, D., and Krings, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *J. Pers. Soc. Psychol.* 81, 116. doi: 10.1093/gmo/9781561592630.article.47256
- Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behav. Brain Sci.* 32, 249–266. doi: 10.1017/S0140525X09990951
- Balkwell, J. W., and Berger, J. (1996). Gender, status, and behavior in task situations. *Soc. Psychol. Q.* 59, 273–283. doi: 10.2307/2787023
- Bassett, R. E., Staton-Spicer, A. Q., and Whitehead, J. L. (1979). Effects of source attire on judgments of credibility. *Cent. States Speech J.* 30, 282–285. doi: 10.1080/10510977909368022
- Betz, L. L. (1986). *Despotism and Differential Reproduction: A Darwinian View of History*. New York, NY: Aldine Publishing Co.
- Binkley, C. (2008). Business casual: all business, never casual. *Wall Street J.* D1–D8.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 52, 345–370. doi: 10.1007/BF02294361
- Brase, G. L., and Guy, E. C. (2004). The demographics of mate value and self-esteem. *Pers. Individ. Differ.* 36, 471–484. doi: 10.1016/S0191-8869(03)00117-X
- Brase, G. L., and Richmond, J. (2004). The white-coat effect: Physician attire and perceived authority, friendliness, and attractiveness. *J. Appl. Soc. Psychol.* 34, 2469–2481.
- Brooks, R. C., Jordan, A. L., Shelly, J., and Dixon, B. J. (2015). The multivariate evolution of female body shape in an artificial digital ecosystem. *Evol. Hum. Behav.* 36, 351–358. doi: 10.1016/j.evolhumbehav.2015.02.001
- Busetta, G., Fiorillo, F., and Visalli, E. (2013). *Searching for a Job is a Beauty Contest*. Available at: <https://mpira.uni-muenchen.de/49392/2/>
- Buss, D. M., and Schmitt, D. P. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychol. Rev.* 100, 204–232. doi: 10.1037/0033-295X.100.2.204
- Chen, M. K. (2008). *Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?*. New Haven, CT: Yale University.
- Dixon, B. J. (2016). *Masculinity and femininity*. *Encyclopedia of Evolutionary Psychological Science*. Berlin: Springer.
- Dixon, B. J. (2019). Sexual selection and extended phenotypes in humans. *Adapt. Hum. Behav. Physiol.* 5, 103–107. doi: 10.1007/s40750-018-0106-3
- Dixon, B. J., Duncan, M., and Dixon, A. F. (2015). The role of breast size and areolar pigmentation in perceptions of women's sexual attractiveness, reproductive health, sexual maturity, maternal nurturing abilities, and age. *Arch. Sex. Behav.* 44, 1685–1695. doi: 10.1007/s10508-015-0516-2
- Dixon, B. J., Grimshaw, G. M., Linklater, W. L., and Dixon, A. F. (2011). Eye-tracking of men's preferences for waist-to-hip ratio and breast size of women. *Arch. Sex. Behav.* 40, 43–50. doi: 10.1007/s10508-009-9523-5
- Dixon, B. J., Grimshaw, G. M., Ormsby, D. K., and Dixon, A. F. (2014). Eye-tracking women's preferences for men's somatotypes. *Evol. Hum. Behav.* 35, 73–79. doi: 10.1016/j.evolhumbehav.2013.10.003
- Dixon, B. J. W., Rantala, M. J., and Brooks, R. C. (2019). Cross-cultural variation in women's preferences for men's body hair. *Adapt. Hum. Behav. Physiol.* doi: 10.1007/s40750-019-0107-x.
- Dixon, B. J. W., Rantala, M. J., Melo, E. F., and Brooks, R. C. (2017). Beards and the big city: displays of masculinity may be amplified under crowded conditions. *Evol. Hum. Behav.* 38, 259–264. doi: 10.1016/j.evolhumbehav.2016.10.009
- Dixon, B. J. W., Sulikowski, D., Gouda-Vossos, A., Rantala, M. J., and Brooks, R. C. (2016). The masculinity paradox: facial masculinity and beardedness interact to determine women's ratings of men's facial attractiveness. *J. Evol. Biol.* 29, 2311–2320. doi: 10.1111/jeb.12958
- Dunn, M. J., and Doria, M. V. (2010). Simulated attraction increases opposite sex attractiveness ratings in females but not males. *J. Soc. Evol. Cult. Psychol.* 4, 1–17. doi: 10.1037/h0099305
- Dunn, M. J., and Hill, A. (2014). Manipulated luxury-apartment ownership enhances opposite-sex attraction in females but not males. *J. Evol. Psychol.* 12, 1–17. doi: 10.1556/JEP.12.2014.1.1
- Dunn, M. J., and Searle, R. (2010). Effect of manipulated prestige-car ownership on both sex attractiveness ratings. *Br. J. Psychol.* 101, 69–80. doi: 10.1348/000712609X417319
- Eagly, A. H., and Johnson, B. T. (1990). Gender and leadership style: a meta-analysis. *Psychol. Bull.* 108, 233–256. doi: 10.1037/0033-2909.108.2.233
- Eagly, A. H., and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* 109, 573–598. doi: 10.1037/0033-295X.109.3.573
- Forsythe, S., Drake, M. F., and Cox, C. E. (1985). Influence of applicant's dress on interviewer's selection decisions. *J. Appl. Psychol.* 70, 374–378. doi: 10.1037/0021-9010.70.2.374
- Forsythe, S. M. (1990). Effect of applicant's clothing on interviewer's decision to hire. *J. Appl. Soc. Psychol.* 20, 1579–1595. doi: 10.1111/j.1559-1816.1990.tb01494.x
- Forsythe, S. M., Drake, M. F., and Cox, C. A. (1984). Dress as an influence on the perceptions of management characteristics in women. *Fam. Consum. Sci. Res. J.* 13, 112–121. doi: 10.1177/1077727X8401300203
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., and McCormick, C. M. (2015). Evidence from meta analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS One* 10:e0132726. doi: 10.1371/journal.pone.0132726
- Glick, P., Larsen, S., Johnson, C., and Branstetter, H. (2005). Evaluations of sexy women in low-and high-status jobs. *Psychol. Women Q.* 29, 389–395. doi: 10.1111/j.1471-6402.2005.00238.x
- Goktepe, J. R., and Schneier, C. E. (1989). Role of sex, gender roles, and attraction in predicting emergent leaders. *J. Appl. Psychol.* 74, 165–167. doi: 10.1037/0021-9010.74.1.165
- Gouda-Vossos, A., Dixon, B. J., and Brooks, R. C. (2016). Sexual conflict and gender gap effects: associations between social context and sex on rated attractiveness and economic status. *PLoS One* 11:e0146269. doi: 10.1371/journal.pone.0146269
- Gouda-Vossos, A., Nakagawa, S., Dixon, B. J. W., and Brooks, R. C. (2018). Mate choice copying in humans: a systematic review and meta-analysis. *Adapt. Hum. Behav. Physiol.* 4, 364–384. doi: 10.1007/s40750-018-0099-y
- Hamermesh, D. S., and Biddle, J. E. (1993). Beauty and the labor market. *Natl. Bureau Econ. Res.* 84, 1174–1194. doi: 10.3386/w4518
- Hanson, F. I., Olson, J., and Russell, J. (1991). 'Attractiveness and Income for Men and Women. *J. Appl. Soc. Psychol.* 21, 1039–1057. doi: 10.1111/j.1559-1816.1991.tb00458.x
- Heilman, M. E., Wallen, A. S., Fuchs, D., and Tamkins, M. M. (2004). Penalties for success: reactions to women who succeed at male gender-typed tasks. *J. Appl. Psychol.* 89, 416–427. doi: 10.1037/0021-9010.89.3.416
- Hill, A. K., Hunt, J., Welling, L. L., Cárdenas, R. A., Rotella, M. A., Wheatley, J., et al. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evol. Hum. Behav.* 34, 334–341. doi: 10.1111/j.1420-9101.2008.01633.x
- Hill, S. E., and Buss, D. M. (2008). The mere presence of opposite-sex others on judgments of sexual and romantic desirability: opposite effects for men and women. *Pers. Soc. Psychol. Bull.* 34, 635–647. doi: 10.1177/0146167207313728
- Hopcroft, R. L. (2006). Sex, status, and reproductive success in the contemporary United States. *Evol. Hum. Behav.* 27, 104–120. doi: 10.1016/j.evolhumbehav.2005.07.004
- Howlett, N., Pine, K., Orakçioğlu, I., and Fletcher, B. (2013). The influence of clothing on first impressions: rapid and positive responses to minor changes in male attire. *J. Fashion Mark. Manage.* 17, 38–48. doi: 10.1108/13612021311305128
- Howlett, N., Pine, K. J., Cahill, N., Orakçioğlu, I., and Fletcher, B. C. (2015). Unbuttoned: the interaction between provocativeness of female work attire and occupational status. *Sex Roles* 72, 105–116. doi: 10.1007/s11199-015-0450-8
- Janif, J. Z., Brooks, R. C., and Dixon, B. J. (2014). Negative frequency-dependent preferences and variation in male facial hair. *Biol. Lett.* 10:20130958. doi: 10.1098/rsbl.2013.0958
- Janif, J. Z., Brooks, R. C., and Dixon, B. J. (2015). Are preferences for women's hair color frequency dependent? *Adapt. Hum. Behav. Physiol.* 1, 54–71. doi: 10.1007/s40750-014-0008-y
- Jasińska, G., Ziolkiewicz, A., Ellison, P. T., Lipson, S. F., and Thune, I. (2004). Large breasts and narrow waists indicate high reproductive potential in women. *Proc. R. Soc. B Biol. Sci.* 271, 1213–1217. doi: 10.1098/rspb.2004.2712

- Kinsey, A. C., Pomeroy, W. B., Martin, C. E., and Gebhard, P. H. (1953). *Sexual Behavior in the Human Female*. Bloomington, IN: Indiana University Press.
- Kinsey, A. C., Pomeroy, W. B., Martin, C. E., and Sloan, S. (1948). *Sexual Behavior in the Human Male*, Vol. 1. Philadelphia, PA: Saunders.
- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., and Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum. Behav.* 39, 424–436. doi: 10.1016/j.evolhumbehav.2018.03.008
- Kwantes, C. T., Lin, I. Y., Gidak, N., and Schmidt, K. (2011). The effect of attire on expected occupational outcomes for male employees. *Psychol. Men Masc.* 12, 166–180. doi: 10.1037/a0020872
- Lavin, A., Carr, D., and Davies, T. (2009). Does image matter? Female and male student perceptions of the impact of the female professor's attire. *J. Bus. Account.* 2, 93–106.
- Lavin, A. M., Carr, D. L., and Davies, T. L. (2009). The male professor's attire and student perceptions of instructional quality. *Res. High. Educ. J.* 4:1.
- Li, N. P., Bailey, J. M., Kenrick, D. T., and Linsenmeier, J. A. (2002). The necessities and luxuries of mate preferences: testing the tradeoffs. *J. Pers. Soc. Psychol.* 82, 947–955. doi: 10.1037/0022-3514.82.6.947
- Low, B. S. (1990). Sex, power, and resources: male and female strategies of resource acquisition. *Int. J. Contemp. Soc.* 27, 49–73.
- Low, B. S. (2000). "Sex, wealth, and fertility: Old rules, new environments," in *Adaptation and Human Behavior: An Anthropological Perspective*, eds L. Cronk, N. Chagon, and W. Irons (New York, NY: Aldline), 323–344.
- Luoto, S. (2019). An updated theoretical framework for human sexual selection: From ecology, genetics, and life history to extended phenotypes. *Adapt. Hum. Behav. Physiol.* 5, 48–102.
- Maestripieri, D., Henry, A., and Nickels, N. (2017). Explaining financial and prosocial biases in favor of attractive people: interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behav. Brain Sci.* 40:e19. doi: 10.1017/S0140525X16000340
- Mangel, M. (1990). Resource divisibility, predation and group formation. *Anim. Behav.* 39, 1163–1172. doi: 10.1016/S0003-3472(05)80788-2
- Marcinkowska, U. M., Kozlov, M. V., Cai, H., Contreras-Garduño, J., Dixon, B. J., Oana, G. A., et al. (2014). Cross-cultural variation in men's preference for sexual dimorphism in women's faces. *Biol. Lett.* 10:20130850. doi: 10.1098/rsbl.2013.0850
- Melis, A. P., and Semmann, D. (2010). How is human cooperation different? *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2663–2674. doi: 10.1098/rstb.2010.0157
- Milonoff, M., Nummi, P., Nummi, O., and Pienmunne, E. (2007). Male friends, not female company, make a man more attractive. *Ann. Zool. Fennici* 44, 348–354.
- Molloy, J. T., and Potter, S. (1975). *Dress for Success*. Available at: <https://dressforsuccess.org/>
- Morris, T. L., Gorham, J., Cohen, S. H., and Huffman, D. (1996). Fashion in the classroom: effects of attire on student perceptions of instructors in college classes. *Commun. Educ.* 45, 135–148. doi: 10.1080/03634529609379043
- Pettersson, L. J., Dixon, B. J., Little, A. C., and Vasey, P. L. (2015). Viewing time measures of sexual orientation in Samoan cisgender men who engage in sexual interactions with fa'afafine. *PLoS One* 10:e0116529. doi: 10.1371/journal.pone.0116529
- Pettersson, L. J., Dixon, B. J., Little, A. C., and Vasey, P. L. (2016). Reconsidering male bisexuality: sexual activity role and sexual attraction in Samoan men who engage in sexual interactions with Fa'afafine. *Psychol. Sex. Orientat. Gend. Divers.* 3, 11–26. doi: 10.1037/sgd0000160
- Pettersson, L. J., Dixon, B. J., Little, A. C., and Vasey, P. L. (2018). Viewing time and self-report measures of sexual attraction in Samoan cisgender and transgender androphilic males. *Arch. Sex. Behav.* 47, 2427–2434. doi: 10.1007/s10508-018-1267-7
- Pfann, G. A., Biddle, J. E., Hamermesh, D. S., and Bosman, C. M. (2000). Business success and businesses' beauty capital. *Econ. Lett.* 67, 201–207. doi: 10.1016/S0165-1765(99)00255-4
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Roach-Higgins, M. E., and Eicher, J. B. (1992). Dress and identity. *Cloth. Text. Res. J.* 10, 1–8. doi: 10.1177/0887302X9201000401
- Rosenblat, T. S. (2008). The beauty premium: physical attractiveness and gender in dictator games. *Negotiat. J.* 24, 465–481. doi: 10.1111/j.1571-9979.2008.00198.x
- Schmid Mast, M. (2004). Men are hierarchical, women are egalitarian: an implicit gender stereotype. *Swiss J. Psychol.* 63, 107–111. doi: 10.1024/1421-0185.63.2.107
- Schopler, J., Insko, C. A., Wieselquist, J., Pemberton, M., Witcher, B., Kozar, R., et al. (2001). When groups are more competitive than individuals: the domain of the discontinuity effect. *J. Pers. Soc. Psychol.* 80, 632–644. doi: 10.1037/0022-3514.80.4.632
- Shuler, G. A., and McCord, D. M. (2010). Determinants of male attractiveness: "Hotness" ratings as a function of perceived resources. *Am. J. Psychol. Res.* 6, 10–23.
- Singh, D., Dixon, B. J., Jessop, T. S., Morgan, B., and Dixon, A. F. (2010). Cross-cultural consensus for waist-hip ratio and women's attractiveness. *Evol. Hum. Behav.* 31, 176–181. doi: 10.1016/j.evolhumbehav.2009.09.001
- Smith, E. A. (2004). Why do good hunters have higher reproductive success? *Hum. Nat.* 15, 343–364. doi: 10.1007/s12110-004-1013-9
- Townsend, J. M., and Levy, G. D. (1990). Effects of potential partners' physical attractiveness and socioeconomic status on sexuality and partner selection. *Arch. Sex. Behav.* 19, 149–164. doi: 10.1007/BF01542229
- Valentova, J. V., Varella, M., Bártová, K., Štirbová, Z., and Dixon, B. J. W. (2017). Mate preferences and choices for facial and body hair in heterosexual women and homosexual men: effects of sex, population, homogamy, and imprinting-like effects. *Evol. Hum. Behav.* 38, 241–248. doi: 10.1016/j.evolhumbehav.2016.10.007
- Von Rueden, C., Gurven, M., and Kaplan, H. (2010). Why do men seek status? Fitness payoffs to dominance and prestige. *Proc. R. Soc. Lond. B Biol. Sci.* 278, 2223–2232. doi: 10.1098/rspb.2010.2145
- Von Rueden, C. R. (2011). *The Acquisition of Social Status by Males in Small-Scale Human Societies (with an Emphasis on the Tsimane of Bolivia)*. Santa Barbara, CA: University of California.
- Von Rueden, C. R., and Jaeggi, A. V. (2016). Men's status and reproductive success in 33 nonindustrial societies: effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10824–10829. doi: 10.1073/pnas.1606800113
- Vugt, M. V., Cremer, D. D., and Janssen, D. P. (2007). Gender differences in cooperation and competition: the male-warrior hypothesis. *Psychol. Sci.* 18, 19–23. doi: 10.1111/j.1467-9280.2007.01842.x
- Wang, Y., and Griskevicius, V. (2013). Conspicuous consumption, relationships, and rivals: women's luxury products as signals to other women. *J. Consum. Res.* 40, 834–854. doi: 10.1086/673256
- Waynforth, D. (2007). Mate choice copying in humans. *Hum. Nat.* 18, 264–271. doi: 10.1007/s12110-007-9004-2
- Williams, P. M. (2002). Community strategies: mainstreaming sustainable development and strategic planning? *Sustain. Dev.* 10, 197–205. doi: 10.1002/sd.197

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Gouda-Vossos, Brooks and Dixon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Roar of a Champion: Loudness and Voice Pitch Predict Perceived Fighting Ability but Not Success in MMA Fighters

Pavel Šebesta^{1,2*}, Vít Třebický^{1,3}, Jitka Fialová^{1,3} and Jan Havlíček^{1,3*}

¹ National Institute of Mental Health, Klecany, Czechia, ² Faculty of Humanities, Charles University, Prague, Czechia, ³ Faculty of Science, Charles University, Prague, Czechia

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Benedict C. Jones,
University of Glasgow,
United Kingdom
Phil McAleer,
University of Glasgow,
United Kingdom

*Correspondence:

Jan Havlíček
jhavlicek@natur.cuni.cz
Pavel Šebesta
pavelsebest@email.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 12 October 2018

Accepted: 01 April 2019

Published: 30 April 2019

Citation:

Šebesta P, Třebický V, Fialová J and
Havlíček J (2019) Roar of a Champion:
Loudness and Voice Pitch Predict
Perceived Fighting Ability but Not
Success in MMA Fighters.
Front. Psychol. 10:859.
doi: 10.3389/fpsyg.2019.00859

Historically, antagonistic interactions have been a crucial determinant of access to various fitness-affecting resources. In many vertebrate species, information about relative fighting ability is conveyed, among other things, by vocalization. Previous research found that men's upper-body strength can be assessed from voice. In the present study, we tested formidability perception of intimidating vocalization (roars) and a short utterance produced by amateur male MMA fighters attending the amateur European Championships in relation to their physical fitness indicators and fighting success. We also tested acoustic predictors of the perceived formidability. We found that body height, weight, and physical fitness failed to predict perceived formidability either from speech or from the roars. Similarly, there was no significant association between formidability of the roars and utterances and actual fighting success. Perceived formidability was predicted mainly by roars' and utterances' intensity and roars' harmonics-to-noise ratio and duration. Interestingly, fundamental frequency (F_0) predicted formidability ratings in both roars and utterances but in an opposite manner, so that low F_0 utterances but high F_0 roars were rated as more formidable. Our results suggest that formidability perception is primarily driven by intensity and duration of the vocalizations.

Keywords: speech, roar, vocalization, handgrip, competition, perception, human

INTRODUCTION

Historical and ethnographic evidence shows that physical encounters were a frequent way of resolving conflicts (Manson et al., 1991; Keeley, 1997). Cross-culturally, man's fighting ability is a powerful determinant of access to resources (Daly and Wilson, 1988). These findings are complemented by psychological studies which show that stronger men are more prone to anger (Archer and Thanzami, 2007; Sell et al., 2009b). One may therefore expect that cognitive processes evolved for assessing the threat potential of a prospective opponent (Sell et al., 2009a; Puts, 2010). Earlier research tended to focus on visual cues to the threat potential. It has been demonstrated, for instance, that people can relatively accurately assess physical strength from images of body and face (Sell et al., 2009a; Holzleitner and Perrett, 2016; Kordsmeyer et al., 2018). Moreover, it seems that based on facial images raters can predict winners of mixed martial arts (MMA) fights (Třebický et al., 2013; Little et al., 2015; but see Třebický et al., 2019).

The cues to threat potential are not restricted to the visual modality but evidence regarding vocal indicators of threat potential is rather mixed. On one hand, it was reported that both men and

women can accurately assess men's physical strength from voice irrespectively of the language used (Sell et al., 2010). On the other hand, fighting ability assessed by acquaintances did not correlate with ratings of fighting ability based on vocal stimuli (Doll et al., 2014). Han et al. (2017) likewise reported no association between a composite measure of threat potential, consisting of handgrip strength, body height and weight, and the perceived vocal threat potential.

Importantly, all of the abovementioned studies used speech as their acoustic stimuli. Humans, however, produce also various other vocalizations, such as laughter, roars, screams and grunts, and these so far received only limited attention. This contrasts with evidence from a number of vertebrate species, including primates, which shows that vocal displays are frequently part of male intrasexual competition (Bradbury and Vehrencamp, 2011) and can indicate fighting ability (for evidence in red deer, see Clutton-Brock and Albon, 1979; for baboons, see Kitchen et al., 2003). In humans, it has recently been shown that tennis players who produce grunts with a lower fundamental frequency (F_0) are more likely to win and listeners can to some extent predict match outcome from the grunts (Raine et al., 2017). Similarly, Raine et al. (2018a) reported that listeners accurately assess relative strength and body height from aggressive roars in both men and women.

In our complementary study, we tested predictors of perceived formidability using acoustic cues. It ought to be noted, however, that Raine et al. (2018a) and the current study differ in several important respects. First of all, Raine et al. focused on two important components of threat potential (height and strength), but threat potential and/or *perceived* formidability undoubtedly include other components as well. These may include morphological characteristics, such as body weight and lean muscle mass, as well as physical abilities other than isometric strength, for instance respiratory fitness. Secondly, while one can expect that threat potential is a predictor of outcomes of real-life fights, it cannot be entirely equated with fighting success.

To address these questions, we recorded both verbal and non-verbal vocalizations (utterances and roars) of amateur male MMA athletes along with (i) measurements of their body composition, isometric strength, and spirometry, and collected data regarding their (ii) fighting success.

We hypothesized that formidability perceived from vocalization should correlate with height, weight, and muscle mass as well as physical fitness indicators, such as strength and lung capacity. We also predicted that perceived formidability is positively associated with fighting success. Further, we performed an acoustic analysis to identify which parameters predict the perception of formidability from both roars and utterances. We hypothesized that perceived formidability is related to the F_0 and intensity in both verbal and non-verbal vocalizations.

MATERIALS AND METHODS

All procedures applied in this study were in accordance with ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration. The study

was approved by the Institutional Review Board of the National Institute of Mental Health, Czech Republic (Ref. num. 28/15). All target participants were provided with a brief description of the study and approved their participation by signing informed consent. The present study is part of a larger project investigating multi-modal perception of traits associated with sexual selection and characteristics related to competition outcome.

Targets

Data collection took place during 2016 IMMAF European Open Championships of Amateur MMA held in Prague (Czech Republic), which hosted a total of 155 contestants (incl. 20 women) from 30 countries (based on data from MyNextMatch.com). Contestants were approached by researchers during registration on site, 1 day before the start of the tournament. We focused on male athletes because championship attendance was highly biased toward male athletes and we thus managed to collect data from only three female athletes.

Forty male amateur MMA fighters (mean age = 24, $SD = 4.4$, range = 19–33 years), naïve to our project's aims, participated in the study. To assess a possible effect of weight category, we merged the weight categories used by competition organizers (Flyweight, $N = 1$; Bantamweight, $N = 7$; Featherweight, $N = 4$; Lightweight, $N = 4$; Welterweight, $N = 7$; Middleweight, $N = 5$; Light Heavyweight, $N = 5$; Heavyweight, $N = 4$; and Super Heavyweight, $N = 3$) into just three categories: Lightweight ($N = 12$; consisting of Flyweight, Bantamweight, and Featherweight categories), Middleweight ($N = 16$; consisting of Lightweight, Welterweight, and Middleweight categories) and Heavyweight ($N = 12$; consisting of Light Heavyweight, Heavyweight, and Super Heavyweight) following procedure in Třebický et al. (2013). All targets reported their basic demographics, age, and total fighting record, from which computed their fighting success as a proportion of the number of wins relative to the total number of fights. Fighting success was calculated only for fighters whose record included more than two fights. Analyses involving fighting success are therefore based on 29 individuals. For technical reasons, we managed to obtain lung capacity measures from 34 individuals. For descriptive statistics, see Table 1. All other analyses are based on the complete dataset of 40 individuals. Participants in the study were financially reimbursed with 400 CZK (app. €15) and verbally debriefed upon completing their participation.

Body Measurements

Body height was measured by Vít Třebický using anthropometer Trystom A-213. Participants were standing with their back against a wall, looking directly ahead, and body height was measured from Vertex to ground to a nearest millimeter (Hall et al., 2007).

Body weight, amount of body fat, and muscle mass were measured by Jitka Fialová (JF) using bio impedance Tanita MC-980 scale (Athlete setting; Vaara et al., 2012). Testing was performed in a standing position while standing on and holding in hands the measuring electrodes with arms hanging freely along

TABLE 1 | Target descriptive statistics.

Characteristic	N	Mean	SD	Minimum	Maximum
Age (yrs)	40	24	4.397	19	33
Score (wins/fights ratio)	29	0.682	0.208	0	1
Body Height (cm)	40	176.7	8.687	161.2	196
Body Weight (kg)	40	79.345	20.593	56.5	134.1
Body Fat (kg)	40	8.047	9.077	1.7	40.1
Muscle Mass (kg)	40	67.775	12.254	50.3	97.6
Maximal Handgrip Strength (right) (kp)	40	52.225	10.34	34.2	76
Maximal Handgrip Strength (left) (kp)	40	51.625	9.127	34.3	68
Mean Handgrip Strength (kp)	40	49.243	8.939	33.283	63.817
FVC (l)	34	4.515	0.879	2.46	6.55
FEV ₁ (l)	34	4.296	0.747	2.45	5.77
PEF (l)	34	8.469	2.471	3.86	11.73
Roar F ₀ (Hz)	40	392.85	56.893	243.503	492.725
Roar F ₃ (Hz)	40	2624.125	233.835	2196	3150
Roar HNR (dB)	40	9.169	5.068	2.926	24.786
Roar Intensity (dB)*	40	54.599	4.68	44.619	61.819
Roar Duration (s)	40	0.641	0.312	0.138	1.280
Utterance F ₀ (Hz)	40	127.771	18.387	100.844	179.099
Utterance F ₃ (Hz)	40	2546.02	98.66	2310.286	2727.487
Utterance HNR (dB)	40	8.929	2.88	3.452	14.205
Utterance Intensity (dB)*	40	13.721	3.803	6.643	23.057
Utterance Duration (s)	40	8.024	2.81	2.898	19.41

*Intensity values are reported as analyzed from the original recordings (i.e., before being post-processed).

the body. Participants were wearing underwear only (Pinilla et al., 1992).

Physical Fitness Measurements

Handgrip strength was measured by JF using Takei TKK 5401 digital hand dynamometer (Vidal Andreato et al., 2011; Bonitch-Góngora et al., 2013). While undergoing the handgrip test, the athletes were instructed to stand straight with arms alongside their body. They had 3 attempts with each hand, alternated hands between attempts, and we used the “best test” method, meaning the attempt with the highest value of handgrip strength for each hand was recorded. Maximal handgrip strength between left and right hand was closely correlated ($r = 0.808$ 95% CI [0.664, 0.894], $p < 0.001$, $N = 40$) and paired sample t -test showed no statistical difference between the maximal strength of left and right hand [$t_{(39)} = 0.618$, $p = 0.54$, mean difference = 0.6 kp]. In all further analyses involving handgrip strength, we therefore represent handgrip strength by the mean of both hands “best test” score.

Measures of lung capacity were taken by JF using MicroLab ML3500 MK8. Three standing forced vital capacity (FVC) maneuvers were performed, “best test” method applied, and we recorded the maneuver with the highest recorded FVC value along with Forced expiratory volume in the first second (FEV₁) and Peak expiratory flow (PEF). The “best test” method is a widely used and recommended approach in research employing spirometry (Crapo et al., 1981; Havryk et al., 2002). FVC is the maximal volume of air exhaled with maximally forced effort from maximal inspiration delivered during expiration made as

forcefully and completely as possible. In other words, it is vital capacity performed with a maximally forced expiratory effort. FEV₁ is the maximal volume of air exhaled in the first second of forced expiration from a position of full inspiration, and PEF represents the maximum expiratory flow achieved by maximum forced expiration from the point of maximal lung inflation (Miller et al., 2005).

Vocal Stimuli Recordings and Processing

Acoustic stimuli were recorded by Pavel Šebesta (PŠ) using Sony PMC-D90 portable audio recorder (in-built microphone sensitivity 20–40 kHz). Recorder was equipped with a windscreens (AD-PCM1), mounted on a tripod with acoustic reflection shield and placed in a portable, acoustically treated booth to reduce any potential echoes and ambient noises. Recordings were captured at 24 bit/96 kHz in WAV format. Participants stood 1.5 meters from the recorder and Levels setting was kept constant in the course of all recordings to standardize recording intensity and to prevent clipping.

Participants were instructed to count from 1 to 10 in their native language and then perform three intimidating roars (their instruction was: “Roar three times, as much as you can, to intimidate a potential opponent”). For ratings and analyses, we use only the second roar because the first might be affected by the novelty of the task and the third by a potential decrease of effort (for differences between the three roars, see Supplementary Material **Tables S1–S11**). For examples of roars see **Audios S1, S2** and for utterances see **Audios S3, S4**.

Subsequent processing and acoustical analysis were performed by PŠ in Audacity 2.1.3 (Audacity Team, 2018) and Praat 5.4.09 (Boersma and Weenink, 2015). Roars and utterance levels were increased by +20 dB and +35 dB, respectively, while interindividual variation in vocalization intensity remained unmodified. This intensity adjustment was necessary because most utterance recordings were not sufficiently loud even at the highest volume settings. The employed adjustments in roars was the highest possible that did not introduce clipping in any of the recordings. We measured the mean intensity and duration of volume-adjusted utterances and roars. Mean F_0 was measured by autocorrelation method. Preset parameters for F_0 extraction were used, with a 75 Hz pitch floor in accordance with Praat programmers' recommendations and 300 Hz pitch ceiling based on a visual inspection of spectrographs (for similar approach see also Šebesta et al., 2017). The 300 Hz pitch ceiling recommended for utterances was not suitable for the roars. We visually inspected Praat's pitch contours in the Editing window. Most roar recordings showed erroneous F_0 measurements (see **Figure S1** for an example), which rendered the standard Praat's F_0 extraction method unreliable for this type of acoustic stimuli (for similar issues with F_0 extraction, see Raine et al., 2017). F_0 tracking frequently failed in the middle of recording or even unexpectedly "jumped down." This is possibly due to chaotic and subharmonic phenomena found in roars (Fitch et al., 2002). For this reason, we decided to use, as a F_0 analog, the long-term averaged Fast Fourier transformed (FFT) spectral peak frequency (see **Figure S2** for an example), corresponding to the first harmonic (verified by a visual inspection of harmonic structure). Further, we used standard Praat methods for harmonics-to-noise ratio (HNR; autocorrelation method, preset parameters) measurements for whole utterance recordings, and one second long snips from the initial part of the roars close to the spectrogram plateau where Praat's autocorrelation algorithm was able to track F_0 . Mean formant levels in speech (F_1 – F_4) were measured by Burg method. In roars, however, only a peak around 2–3 kHz (which is in expected range for the third formant) was apparent by a visual inspection of long-term average spectrums (LTAS) and clearly distinguishable from other harmonics. Audacity's "Plot spectrum" feature (Spectrum, 1,024 window size, Hanning window) was used for the 2–3 kHz peak measurement. Because we were able to reliably extract only the third formant (F_3) from roars and the first and second formants in speech are highly affected by speech content, we decided to use in subsequent analyses only the third formant of both utterances and roars to enable comparison.

Rating Sessions

In total, 31 men (mean age = 27.1, SD = 5.2, range = 20–36 years) and 32 women (mean age = 24.4, SD = 4.3, range = 18–33 years), mainly students at the Charles University, Prague, Czech Republic, took part in rating sessions.

Raters were recruited via social media advertisements and mailing lists of participants from previous studies. After completing participation, they were financially reimbursed with 100 CZK (~€4), a small snack, and received a debriefing leaflet about the purpose of the study.

Raters were asked to assess the formidability ("Jak moc by byl tento muž úspěšný, kdyby se dostal do fyzického souboje?" / "How successful would this man be if he was involved in a physical confrontation?") of a given recording on a 7-point verbally anchored scale (from "1–velice neúspěšný" / "not successful at all," to "7–velice úspěšný" / "highly successful"). Each participant rated all roar and utterance stimuli. To reduce participant fatigue, the rating was divided in two sessions 1 week apart. In the first session, participants rated half of the set of all roars and utterances in a randomized order. Individual stimuli within the set were randomized as well. In the second session, participants rated the remaining half of the stimuli in the same fashion.

Ratings took place in a quiet perception lab room with negligible ambient sounds. Focusrite Scarlett Solo Gen 2 audio I/O interface (22 Hz–22 kHz RCA output) and two Yamaha HS-7 active reference studio monitor speakers (43 Hz–30 kHz @ 95W, LF 60 W, HF 35 W output) were used to present stimuli in WAV format. Raters were seated 2.8 meters in front of and in focus of the speakers. We opted for speakers, instead of commonly used headphones, because it is a more ecologically valid approach to presenting stimuli in terms of sonic characteristics of roaring. Loudness of the playback was kept standard during the presentation, with the loudest roar registering 87 dB (measured with OnePlus One smartphone and Smart Tools® Sound meter 1.6.12 app). This is a level which, all authors agreed, was very naturalistic but not overwhelmingly loud.

Statistical Analyses

All statistical tests were performed in JASP 0.9.0.1 (JASP Team, 2018) and jamovi 0.9.1.7 (jamovi project, 2018). McDonald's ω statistics was used to estimate interrater agreement (Dunn et al., 2014). To test for potential sex differences in ratings, a paired samples t -test was carried out. Associations between ratings by men and women were tested by bivariate correlations using Pearson's r coefficient with 95% CIs [lower limit, upper limit]. Potential differences between the maximal strength of left and right hand were tested with paired samples t -test, and associations between the left and right hand strength were tested by bivariate correlations using Pearson's r coefficient with 95% CIs. Cohen's d , as an effect size measure, was used for means comparisons. To assess the relative contribution of performance-related and acoustic measures to the perceived roar and utterance formidability, we performed Linear mixed effects model (using REML fit) with individual rater ID and target stimuli ID as random intercepts. This approach accounted for variation on the level of individual raters and for variation on the level of individual stimuli. It also accounted for potential bias due to the data aggregation. To assess acoustic predictors of fighting success, we ran a linear regression analysis (Enter method). As measures of variability explained by regression, we list model R^2 values, while standardized β s and their 95% CI are reported for entered coefficients.

Data Availability

Datasets generated and analyzed during the current study are available in the **Supplementary Material** of this article (Tables S20, S21).

RESULTS

Sex Differences in Perceived Formidability Utterances

McDonald's ω scores of male ($\omega = 0.954$) and female ($\omega = 0.933$) ratings of formidability of utterances showed a high interrater agreement. We have therefore used mean formidability ratings given to the individual utterances separately by male and female raters. Perceived formidability of utterances was likewise highly correlated between men and women ($r = 0.93$ 95% CI [0.871, 0.963], $p < 0.001$, $N = 40$). Paired sample t -test showed a statistically significant sex difference in formidability ratings with men giving higher ratings [$t_{(39)} = 9.165$, $p < 0.001$, Cohen's $d = 1.449$, mean difference = 0.368] (for descriptive statistics, see Table 2). Although mean ratings of utterance formidability differed between sexes, all further analyses are reported with ratings combined because the results are virtually the same when analyzed separately. For results based on female and male ratings separately, see Supplementary Material Tables S12–S19.

Roars

McDonald's ω scores of males ($\omega = 0.953$) and females ($\omega = 0.924$) ratings of roar formidability showed a high interrater agreement. In subsequent analyses, we have therefore used mean formidability ratings given to the individual roars separately by male and female raters. Further, we found a high correlation between roar formidability ratings assigned by men and by women ($r = 0.973$ 95% CI [0.95, 0.986], $p < 0.001$, $N = 40$). Paired sample t -test showed statistically significant difference between the sexes in roar formidability ratings with women giving higher ratings [$t_{(39)} = 2.695$, $p = 0.045$, Cohen's $d = 0.426$, mean difference = 0.132]. For descriptive statistics, see Table 2.

Formidability of Utterances and Roars as a Predictor of Fighting Success

To test whether formidability perception from roars and utterances predicts fighting success, we ran bivariate Pearson's correlations. We found that neither in utterances ($r = -0.045$ 95% CI [-0.405, 0.327], $p = 0.817$, $N = 29$) nor in roars ($r = -0.115$ 95% CI [-0.462, 0.263], $p = 0.554$, $N = 29$) was formidability perception associated with actual fighting success. To explore whether the effect is modulated by the weight categories, we grouped the fighters in three weight categories (lightweight, middleweight, and heavyweight) and entered this variable into the linear regression. Even after this modification, however, the overall model was not formally significant either in utterances [$F_{(3, 25)} = 1.841$, $p = 0.166$, $R^2 = 0.181$] or in roars [$F_{(3, 25)} = 0.683$, $p = 0.571$, $R^2 = 0.076$].

Physical Fitness Predictors of Perceived Formidability

First, we ran exploratory correlational analyses to assess relationships between the physical fitness variables (see Supplementary Material Table S22). Body weight, muscle mass, and fat mass were all highly positively intercorrelated ($r_s > 0.757$, $p_s < 0.001$, $N = 40$). To avoid collinearity and to facilitate interpretation of the findings, we used only body weight

TABLE 2 | Formidability rating descriptive statistics.

Stimuli	Raters	Rating		McDonald's ω
		Mean	SD	
Utterance	Total	3.404	0.663	0.947
	Women	3.22	0.687	0.933
	Men	3.588	0.664	0.954
Roar	Total	3.92	1.15	0.939
	Women	3.986	1.237	0.924
	Men	3.854	1.079	0.953

TABLE 3 | Summary of linear mixed effects model analysis for physical fitness predictors of perceived fighting ability based on utterances and roars.

	Predictors	Estimate	t	p	95% CI
Utterances	Age (yrs)	0.024	0.923	0.364	-0.027, 0.074
	Height (cm)	0.050	1.943	0.062	-0.00004, 0.010
	Weight (kg)	-0.010	-1.039	0.308	-0.028, 0.009
	FVC (l)	-0.277	-1.953	0.061	-0.556, 0.0001
	PEF (l)	0.016	0.323	0.749	-0.081, 0.113
	Handgrip strength (kp)	0.011	0.735	0.469	-0.019, 0.041
Roars	Age (yrs)	0.073	1.505	0.144	-0.022, 0.169
	Height (cm)	-0.044	-0.899	0.377	-0.138, 0.051
	Weight (kg)	0.017	0.971	0.340	-0.017, 0.052
	FVC (l)	0.402	1.496	0.146	-0.125, 0.929
	PEF (l)	-0.114	-1.223	0.232	-0.297, 0.069
	Handgrip strength (kp)	0.004	0.153	0.880	-0.052, 0.061

in the subsequent analyses. FVC and FEV₁ spirometry measures were likewise highly positively correlated ($r = 0.935$ 95% CI [0.872, 0.967], $p < 0.001$, $N = 34$), which is why we decided to omit the FEV₁ from subsequent analyses.

Linear mixed model analyses were run with age, height, weight, FVC, PEF, and handgrip strength as fixed effect predictors to assess whether physical fitness parameters predict the perceived formidability of utterances and roars. The overall model for utterances explained 44.9% of variance (R^2 conditional) and fixed factors explained 5.4% of variance (R^2 marginal). None of the physical fitness predictors for the formidability of utterances was formally significant. The overall model for roars explained 60.1% of variance (R^2 conditional), while fixed factors explained 8.2% of variance (R^2 marginal). Similarly, none of the predictors of perceived formidability in roars were significant. For an overview of the results, see Table 3.

Acoustic Predictors of Perceived Formidability

Linear mixed model analyses were run to predict perceived formidability from utterances and roars with F_0 , F_3 , HNR, intensity, and duration entered as independent predictors. For utterances, the overall model explained 44.1% of variance (R^2

TABLE 4 | Summary of linear mixed effects model analysis for acoustic predictors of perceived formidability based on utterances and roars.

	Predictors	Estimate	t	P	95% CI
Utterances	F ₀ (Hz)	−0.017	−3.326	0.002	−0.027, −0.007
	F ₃ (Hz)	−0.0001	−1.213	0.233	−0.003, 0.00006
	HNR (dB)	0.007	0.194	0.847	−0.061, 0.074
	Intensity (dB)	0.140	4.663	< 0.001	0.081, 0.199
	Duration (s)	−0.023	−0.691	0.494	−0.088, 0.042
Roars	F ₀ (Hz)	0.005	3.31	0.002	0.002, 0.008
	F ₃ (Hz)	−0.00004	−1.59	0.121	−0.00009, 0.00001
	HNR (dB)	−0.070	−6.13	< 0.001	−0.093, −0.048
	Intensity (dB)	0.911	4.29	< 0.001	0.495, 1.328
	Duration (s)	0.125	6.96	< 0.001	0.090, 0.160

TABLE 5 | Summary of multiple linear regression analysis for acoustic predictors of fighting success based on utterances and roars.

	Predictors	β	t	p	95% CI
Utterances	F ₀ (Hz)	0.107	0.469	0.644	−0.366, 0.581
	F ₃ (Hz)	−0.076	−0.382	0.706	−0.488, 0.336
	HNR (dB)	−0.036	0.155	0.878	−0.45, 0.522
	Intensity (dB)	0.045	0.181	0.858	−0.469, 0.559
	Duration (s)	−0.338	−1.488	0.150	−0.807, 0.132
Roars	F ₀ (Hz)	0.220	1.476	0.154	−0.185, 1.108
	F ₃ (Hz)	−0.073	−0.318	0.753	−0.472, 0.346
	HNR (dB)	−0.050	−0.093	0.927	−0.420, 0.384
	Intensity (dB)	−0.186	−1.077	0.293	−0.882, 0.278
	Duration (s)	−0.339	−2.023	0.055	−0.921, 0.01

conditional), while fixed factors explained 9.6% of variance (R^2 marginal). We found that F_0 and intensity are significant predictors of perceived formidability. In the case of roars, the overall model explained 57% of variance (R^2 conditional) and fixed factors explained 37.5% of variance (R^2 marginal). We further found that perceived formidability was predicted by the F_0 , HNR, intensity, and duration. For full detail, see **Table 4**.

Acoustic Predictors of Fighting Success

To explore whether any acoustic parameters predict actual MMA fighting success, we ran a multiple linear regression analysis for both utterances and roars. Overall models were not statistically significant in either utterances or roars [Utterances: $F_{(5, 23)} = 0.774$, $p = 0.578$, $R^2 = 0.144$; Roars: $F_{(5, 23)} = 1.107$, $p = 0.384$, $R^2 = 0.194$]. For full results, see **Table 5**.

DISCUSSION

The main goal of this study was to test whether a perception of formidability based on intimidating roars and non-intimidating utterances is related to body parameters such as body height, weight, and to some relevant aspects of physical fitness, such as strength and lung capacity. We have also tested whether

perceived formidability is related to actual fighting success. Finally, we performed an acoustic analysis to investigate which parameters predict perceived formidability and fighting success. In contrast to our predictions, we found that neither body height, weight, or muscle mass predict perceived formidability neither from speech nor roars. We also found no significant association between formidability of the roars and utterances and actual fighting success. Finally, our acoustic analysis showed that the intensity (the acoustic analog of loudness) of both speech and roars is the strongest predictor of perceived formidability. In roars, but not in utterances, lower HNR and longer duration predicted perceived formidability. Moreover, while lower voices (lower F_0) were perceived as more formidable in utterances, the opposite held for the roars.

Our negative findings concerning an association between body height and strength of the roars contrast with results reported in a recent paper by Raine et al. (2018a), where the authors found that the listeners could predict relative body height and handgrip strength from both speech and roars. Such results are further supported by another study which showed a positive association between handgrip strength and perceived strength based on speech (Sell et al., 2010). On the other hand, another two studies found no association between threat potential and perceived fighting ability/dominance from speech (Doll et al., 2014; Han et al., 2017).

There are several possible explanations for such striking differences between our study and results reported in Raine et al. First of all, both Raine et al. (2018a) and Sell et al. (2010) asked participants specifically to assess strength, while our participants rated formidability. Although strength does certainly contribute to overall formidability, there are other important factors which influence it, such as agility or endurance. Moreover, differences in the use of perceptual attributes can, too, affect the association with measures of formidability. Since our main goal was to investigate how people perceive threat potential based on acoustic cues, we used a broader concept of formidability instead of focusing narrowly on the perception of strength. To resolve this issue, future studies should compare ratings of strength and formidability based on acoustic cues and its correlates while employing the same set of stimuli (for results based on the perception of faces and bodies, see Sell et al., 2010).

Secondly, Raine et al. (2018a) in their ratings used an ego-centered approach, i.e., their participants assessed strength relatively to their own strength. We agree that perceivers may be particularly sensitive when it comes to estimating their own chances of winning a confrontation. Nonetheless, several other studies did use absolute ratings, including rating of perception of strength from speech (e.g., Sell et al., 2010), and found positive results. It is possible that even under these conditions, people tend to use the scale relatively to their own prospects. It could also be argued that because our targets were experienced fighters, there should be no difference between the relative and absolute ratings because vast majority of student listeners would rate their formidability as lower than that of MMA fighters in either case. This is supported by a comparison of mean values of handgrip strength between our (**Table 1**) and Raine et al., 2018a (**Supplementary Information**, p. 4) study, although this is only

a very approximate estimate because these two studies used different types of dynamometer and resulting values therefore cannot be directly compared. Alternatively, people might be able to assess formidability irrespective of ego involvement. This is supported by a study which explicitly used the bystander paradigm (Little et al., 2015). In particular, raters were asked to judge from facial photographs who will win a fight and they were successful above the chance level. Once again, to obtain more fine-grained insights into how ego-related context affects the cognitive processes of formidability assessments, future investigations should compare this directly.

Thirdly, in our study we used vocal stimuli from MMA fighters who have extensive experience with physical encounters and some fighters produce roars when winning a fight. It would seem advantageous to employ such a group of participants rather than, for instance, students who are likely to have limited experience with both fighting and roaring. Potential drawback of our sample of fighters may be that because of intense training, they will display little variability in their handgrip strength. Inspection of variation estimates, such as SD, shows that this was not the case (see **Table 1**). The sample size of our stimuli was rather moderate ($N = 40$), but a related study by Raine et al. (2018a) reported positive effect based on smaller sample of the male stimuli.

Finally, one could argue that formidability perception of the roars is related to the effort. This is supported by our acoustic analysis which showed that intensity and duration was the strongest predictor of formidability judgements. It is thus possible that in our sample, motivation and consequently also effort invested in the roars varied among our participants and as a result may have obscured some of the associations with physical characteristics. Alternatively, and perhaps most importantly, the full expression of intimidating roars is not under complete volitional control, which is why it is possible that it can be expressed only in the appropriate context (e.g., when conflict is imminent). Using on-demand roars might not be a problem for judgements of strength but could be a key factor in formidability inferences. Although we acknowledge that this might be a logistically challenging task, the use of real-life non-verbal vocal stimuli which vary little in their motivation and/or effort should thus be preferred. An excellent example in this context is the study by Raine et al. (2017) who used as their stimuli the grunts of professional tennis players.

The acoustic analysis showed that for formidability judgements, intensity and duration are the most salient predictor. This is in agreement with studies on various vertebrate species. For example, male green frogs (*Rana clamitans*) react differently to calls produced by large males as opposed to small ones (Bee et al., 1999; but see Bee, 2002). Similarly, more dominant male baboons produce longer and louder calls (so called “wahoos”) during contest vocalizations (Fischer et al., 2002; Kitchen et al., 2003). Interestingly, many studies on speech perception standardize their vocal stimuli for intensity (because reliable measures of acoustic intensity are logistically difficult to acquire) and therefore cannot assess intensity's contribution to the respective perceptual attribute. However, our results, as well studies on perception of affective states and intentions (Scherer, 1986, 2003; Siegman et al., 1990; Banse and Scherer, 1996), show

that loudness (i.e., the perceptual analog of voice intensity) is an integral and significant part of voice perception. Indeed, the same verbal content expressed in a soft, moderate, or loud voice often has a very different impact on perceivers (Patel et al., 2011). We further found that a low HNR of roars, but not of utterances, is associated with high formidability ratings. Previous studies also show a higher noise in threatening calls than in non-threatening vocalizations and a higher perturbation (lower HNR) in anger vocalization in humans (Patel et al., 2011).

Finally, fundamental frequency was negatively associated with formidability of speech, while associating positively with the formidability of roars. The results of formidability judgements from speech are in agreement with other studies which consistently show that male voices with a lower voice pitch (the perceptual analog of fundamental frequency) are perceived as more dominant and attractive (Puts et al., 2006). This could be a consequence of sex dimorphism in the voice pitch (Rendall et al., 2005; Markova et al., 2016). In contrast, our finding of a positive association between fundamental frequency and formidability judgements of roars came at first as a surprise. On the other hand, one could take into account that high-pitched voices might, similarly to intensity, provide cues about the effort and affective state of the producers, whereby those in a state of high arousal would produce higher F_0 roars. This speculation is supported by studies showing that arousal leads to increase in voice pitch perhaps as a consequence of tension in glottal area (Ekman et al., 1976; van Mersbergen et al., 2017). Moreover, high pitch is in some species associated with threat vocalizations (Stirling, 1971; Portfors, 2007) and in humans, it is associated with anger vocalizations (Scherer and Oshinsky, 1977; Frick, 1986). Fitch et al. (2002) have proposed that subharmonics (portions of F_0) in loud calls are more prevalent and one of the hypothesized effects of this phenomenon is that they perceptually lower the pitch. In other words, a loud vocalization of the same individual that has the same F_0 could sound lower-pitched than if the same vocalization were produced in moderate loudness. Although we were able to detect subharmonics phenomena in a number of high intensity roars in our sample (see Supplementary Material **Table S10**), this effect should be systematically investigated in future studies.

To summarize, we found no significant association between formidability perception of the intimidating roars produced by the MMA fighters and their body height, weight, and physical fitness indicators such as handgrip strength or lung capacity. Neither did we find a correlation between the perceived formidability of their roars and their actual fighting success. This might be because accurate judgements of formidability can be made only on the basis of real-life roars and cannot be reliably performed on demand. It may also be relevant that while roars might be primarily interpreted as intentions (e.g., as affective state of anger), utterances might be interpreted primarily as characteristic of the individual (e.g., as a level of dominance). Alternatively, the association between some acoustic parameters and perceived formidability might be the result of sensory exploitation and have only limited predictive value for actual formidability (Feinberg et al., 2018). We also found that the main acoustic predictors of formidability in roars are intensity, HNR, duration, and to some extent also fundamental frequency.

In a broader context, our study points to a need of further investigations of non-verbal vocalizations in humans. Scholars seem to be so blinded by humans' exceptional gift of speech that they tend to almost completely overlook the fact that this is not our only vocalization. Non-verbal vocalizations are cross-culturally prevalent in human social milieu. This applies not only to preverbal infants (see for instance Lindová et al., 2015) but also to adult humans who produce a wide variety of non-verbal vocalizations in diverse contexts, such as co-laughter, painful injuries, aggressive confrontations, and sexual encounters, to name just few (for some pioneering studies, see Bryant et al., 2016; Raine et al., 2018a,b). We are confident that research into these non-verbal vocal displays will greatly contribute to our understanding of the complexity of human vocal expressions and perhaps also to the evolutionary history of verbal communication in general (Hauser et al., 2002).

AUTHOR CONTRIBUTIONS

PŠ, VT, JF, and JH developed the study concept. Data collection was performed by PŠ, VT, and JF. PŠ performed acoustic analysis of vocal stimuli. VT and PŠ performed data analysis and interpretation jointly with JF and JH. JH, PŠ, and VT drafted the manuscript and JF provided critical revisions. All authors approved the final version of the manuscript for submission.

FUNDING

This research was supported by Czech Science Foundation GAČR P407/16/03899S and by the Ministry of Education, Youth, and Sports (MEYS) NPU I program (No. LO1611) and PROGRES program Q22 at the Faculty of Humanities, Charles University within the Institutional Support for Long-Term Development of Research Organizations from MEYS.

REFERENCES

- Archer, J., and Thanzami, V. (2007). The relation between physical aggression, size and strength, among a sample of young Indian men. *Pers. Individ. Dif.* 43, 627–633. doi: 10.1016/j.paid.2007.01.005
- Audacity Team (2018). *Audacity(R): Free Audio Editor and Recorder (Version 2.3.0)*. Available online at: <https://www.audacityteam.org/download/> (accessed March 8, 2017).
- Banase, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614
- Bee, M. A. (2002). Territorial male bullfrogs (*Rana catesbeiana*) do not assess fighting ability based on size-related variation in acoustic signals. *Behav. Ecol.* 13, 109–124. doi: 10.1093/beheco/13.1.109
- Bee, M. A., Perrill, S. A., and Owen, P. C. (1999). Size assessment in simulated territorial encounters between male green frogs (*Rana clamitans*). *Behav. Ecol. Sociobiol.* 45, 177–184. doi: 10.1007/s002650050551
- Boersma, P., and Weenink, D. (2015). *Praat: Doing phonetics by Computer (Version 5.4.09)*. Available at: <http://www.fon.hum.uva.nl/praat/> (accessed January 20, 2015).
- Bonitch-Góngora, J., Almeida, F., Padial, P., Bonitch-Domínguez, J., and Feriche, B. (2013). Maximal isometric handgrip strength and endurance differences between elite and non-elite young judo athletes. *Arch. Budo.* 9, 239–248.

ACKNOWLEDGMENTS

We would like to thank the International Mixed Martial Arts Federation (IMMAF) and Mixed Martial Arts Association Czech Republic (MMAA) for giving us the opportunity to collect data during the 2016 IMMAF European Open Championships which were held in Prague, Czechia. We are indebted to all the volunteer contestants of the championship and raters for their participation. We wish to thank to Tereza Nevolová, David Stella, and other members of Human Ethology group (www.etologiecloveka.cz) for their help with data collection and ratings, Petr Tureček for help with stimuli randomization, Klára Coufalová, Ph.D. for providing us with physical performance measurements tools and Anna Pilátová, Ph.D. for English proofreading.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00859/full#supplementary-material>

Audio S1 | Sample of highly formidable roar.

Audio S2 | Sample of low formidable roar.

Audio S3 | Sample of highly formidable utterance.

Audio S4 | Sample of low formidable utterance.

Figure S1 | Sample of failed roar F_0 measurement spectrogram.

Figure S2 | Sample of successful roar FFT spectral peak frequency.

Table S1–S19 | Supplementary results.

Table S20 | Dataset ratings.

Table S21 | Dataset targets.

Table S22 | Exploratory correlation table.

- Bradbury, J. W., and Vehrencamp, S. L. (2011). *Principles of Animal Communication*, 2nd ed. Sunderland: Sinauer Associates Inc.
- Bryant, G. A., Fessler, D. M. T., Fusaroli, R., Clint, E., Aaroe, L., Apicella, C. L., et al. (2016). Detecting affiliation in laughter across 24 societies. *Proc. Natl. Acad. Sci. U S A.* 113, 4682–4687. doi: 10.1073/pnas.1524993113
- Clutton-Brock, T. H., and Albon, S. D. (1979). The roaring of red deer and the evolution of honest advertisement. *Behaviour* 69, 145–170.
- Crapo, R. O., Morris, A. T., and Gardner, R. M. (1981). Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am. Rev. Respir. Dis.* 123, 659–664. doi: 10.1164/arrd.1981.123.6.659
- Daly, M., and Wilson, M. I. (1988). *Homicide*. Hawthorne, NY: Aldine.
- Doll, L. M., Hill, A. K., Rotella, M. A., Cárdenas, R. A., Welling, L. L., Wheatley, J. R., et al. (2014). How well do men's faces and voices index mate quality and dominance? *Hum. Nat.* 25, 200–212. doi: 10.1007/s12110-014-9194-3
- Dunn, T. J., Baguley, T., and Brunsden, V. (2014). From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br. J. Psychol.* 105, 399–412. doi: 10.1111/bjop.12046
- Ekman, P., Friesen, W. V., and Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica* 16, 23–27. doi: 10.1515/semi.1976.16.1.23
- Feinberg, D. R., Jones, B. C., and Armstrong, M. M. (2018). Sensory exploitation, sexual dimorphism, and human voice pitch. *Trends Ecol. Evol.* 33, 901–903. doi: 10.1016/j.tree.2018.09.007

- Fischer, J., Hammerschmidt, K., Cheney, D. L., and Seyfarth, R. M. (2002). Acoustic features of male baboon loud calls: Influences of context, age, and individuality. *J. Acoust. Soc. Am.* 111, 1465–1474. doi: 10.1121/1.1433807
- Fitch, W. T., Neubauer, J., and Herzog, H. (2002). Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav.* 63, 407–418. doi: 10.1006/anbe.2001.1912
- Frick, R. W. (1986). The prosodic expression of anger: differentiating threat and frustration. *Aggressive Behav.* 12, 121–128. doi: 10.1002/1098-2337(1986)12:2<121::AID-AB2480120206>3.0.CO;2-F
- Hall, J. G., Allanson, J. E., Gripp, K. W., and Slavotnek, A. M. (2007). *Handbook of Physical Measurements*. New York, NY: Oxford University Press.
- Han, C., Kandrik, M., Hahn, A. C., Fisher, C. I., Feinberg, D. R., Holzleitner, I. J., et al. (2017). Interrelationships among men's threat potential, facial dominance, and vocal dominance. *Evol. Psychol.* 15:1474704917697332. doi: 10.1177/1474704917697332
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- Havryk, A. P., Gilbert, M., and Burgess, K. R. (2002). Spirometry values in Himalayan high altitude residents (Sherpas). *Respir. Physiol. Neurobiol.* 132, 223–232. doi: 10.1016/S1569-9048(02)00072-1
- Holzleitner, I. J., and Perrett, D. I. (2016). Perception of strength from 3D faces is linked to facial cues of physique. *Evol. Hum. Behav.* 37, 217–229. doi: 10.1016/j.evolhumbehav.2015.11.004
- jamovi project (2018). *jamovi (Version 0.9)*. Available online at: <https://www.jamovi.org> (accessed July 15, 2018).
- JASP Team (2018). *JASP (Version 0.9.0.1)*. Available online at: <https://jasp-stats.org/> (accessed July 1, 2018).
- Keeley, L. H. (1997). *War Before Civilization*. New York, NY: Oxford University Press.
- Kitchen, D. M., Seyfarth, R. M., Fischer, J., and Cheney, D. L. (2003). Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behav. Ecol. Sociobiol.* 53, 374–384. doi: 10.1007/s00265-003-0588-1
- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., and Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum. Behav.* 39, 424–436. doi: 10.1016/j.evolhumbehav.2018.03.008
- Lindová, J., Špinka, M., and Nováková, L. (2015). Decoding of baby calls: can adult humans identify the eliciting situation from emotional vocalizations of preverbal infants? *PLoS ONE* 10:e0124317. doi: 10.1371/journal.pone.0124317
- Little, A. C., Třebický, V., Havlíček, J., Roberts, S. C., and Kleisner, K. (2015). Human perception of fighting ability: facial cues predict winners and losers in mixed martial arts fights. *Behav. Ecol.* 26, 1470–1475. doi: 10.1093/beheco/arv089
- Manson, J. H., Wrangham, R. W., Boone, J. L., Chapais, B., Dunbar, R. I. M., Ember, C. R., et al. (1991). Intergroup aggression in chimpanzees and humans. *Curr. Anthropol.* 32, 369–390.
- Markova, D., Richer, L., Pangelinan, M., Schwartz, D. H., Leonard, G., Perron, M., et al. (2016). Age- and sex-related variations in vocal-tract morphology and voice acoustics during adolescence. *Horm. Behav.* 81, 84–96. doi: 10.1016/j.yhbeh.2016.03.001
- Miller, M. R., Hankinson, J., Brusasco, V., Burgos, F., Casaburi, R., Coates, A., et al. (2005). Standardisation of spirometry. *Eur. Respir. J.* 26, 319–338. doi: 10.1183/09031936.05.00034805
- Patel, S., Scherer, K. R., Björkner, E., and Sundberg, J. (2011). Mapping emotions into acoustic space: the role of voice production. *Biol. Psychol.* 87, 93–98. doi: 10.1016/j.biopsycho.2011.02.010
- Pinilla, J. C., Webster, B., Baetz, M., Reeder, B., Hattori, S., and Liu, L. (1992). Effect of body positions and splints in bioelectrical impedance analysis. *JPEN-Parenter. Enter.* 16, 408–412. doi: 10.1177/0148607192016005408
- Portfors, C. V. (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J. Am. Assoc. Lab. Anim.* 46, 28–34.
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Puts, D. A., Gaulin, S. J. C., and Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* 27, 283–296. doi: 10.1016/j.evolhumbehav.2005.11.003
- Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J., and Reby, D. (2018a). Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* 4, 273–280. doi: 10.1016/j.isci.2018.05.002
- Raine, J., Pisanski, K., and Reby, D. (2017). Tennis grunts communicate acoustic cues to sex and contest outcome. *Anim. Behav.* 130, 47–55. doi: 10.1016/j.anbehav.2017.06.022
- Raine, J., Pisanski, K., Simner, J., and Reby, D. (2018b). Vocal communication of simulated pain. *Bioacoustics* 1–23. doi: 10.1080/09524622.2018.1463295
- Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *J. Acoust. Soc. Am.* 117, 944–955. doi: 10.1121/1.1848011
- Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* 99, 143–165. doi: 10.1037/0033-2909.99.2.143
- Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Commun.* 40, 227–256. doi: 10.1016/S0167-6393(02)00084-5
- Scherer, K. R., and Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motiv. Emotion* 1, 331–346. doi: 10.1007/BF00992539
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., et al. (2017). Voices of Africa: acoustic predictors of human male vocal attractiveness. *Anim. Behav.* 127, 205–211. doi: 10.1016/j.anbehav.2017.03.014
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. B Biol. Sci.* 277, 3509–3518. doi: 10.1098/rspb.2010.0769
- Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., and Gurven, M. (2009a). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proc. R. Soc. B Biol. Sci.* 276, 575–584. doi: 10.1098/rspb.2008.1177
- Sell, A., Tooby, J., and Cosmides, L. (2009b). Formidability and the logic of human anger. *Proc. Natl. Acad. Sci. U. S. A.* 106, 15073–15078. doi: 10.1073/pnas.0904312106
- Siegmán, A. W., Anderson, R. A., and Berger, T. (1990). The angry voice: Its effects on the experience of anger and cardiovascular reactivity. *Psychosom. Med.* 52, 631–643. doi: 10.1097/00006842-199011000-00005
- Stirling, I. (1971). Studies on the behaviour of the south Australian fur seal, *Arctocephalus forsteri*. *Aust. J. Zool.* 19, 243–266. doi: 10.1071/ZO9710243
- Třebický, V., Fialová, J., Stella, D., Coufalová, K., Pavelka, R., Kleisner, K., et al. (2019). Predictors of fighting ability inferences based on faces. *Front. Psychol.* 9:2740. doi: 10.3389/fpsyg.2018.02740
- Třebický, V., Havlíček, J., Roberts, S. C., Little, A. C., and Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. *Psychol. Sci.* 24, 1664–1672. doi: 10.1177/0956797613477117
- Vaara, J. P., Kyröläinen, H., Niemi, J., Ohrankämmen, O., Häkkinen, A., Kocay, S., et al. (2012). Associations of maximal strength and muscular endurance test scores with cardiorespiratory fitness and body composition. *J. Strength Cond. Res.* 26, 2078–2086. doi: 10.1519/JSC.0b013e31823b06ff
- van Mersbergen, M., Lyons, P., and Riegler, D. (2017). Vocal responses in heightened states of arousal. *J. Voice* 31, 127.e13–127.e19. doi: 10.1016/j.jvoice.2015.12.011
- Vidal Andreato, L., Franzói de Moraes, S. M., Lopes de Moraes Gomes, T., Del Conti Esteves, J. V., Vidal Andreato, T., and Franchini, E. (2011). Estimated aerobic power, muscular strength and flexibility in elite Brazilian Jiu-Jitsu athletes. *Sci. Sports* 26, 329–337. doi: 10.1016/j.scispo.2010.12.015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Šebesta, Třebický, Fialová and Havlíček. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Evolutionary Theories and Men's Preferences for Women's Waist-to-Hip Ratio: Which Hypotheses Remain? A Systematic Review

Jeanne Bovet*

Stony Brook University, Stony Brook, NY, United States

OPEN ACCESS

Edited by:

Alex L. Jones,
Swansea University, United Kingdom

Reviewed by:

Gayle Brewer,
*University of Liverpool,
United Kingdom*
Lynda Boothroyd,
Durham University, United Kingdom
Robert C. Brooks,
*University of New South Wales,
Australia*

*Correspondence:

Jeanne Bovet
jeanne.bovet@gmail.com

Specialty section:

*This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology*

Received: 08 November 2018

Accepted: 08 May 2019

Published: 04 June 2019

Citation:

Bovet J (2019) Evolutionary Theories and Men's Preferences for Women's Waist-to-Hip Ratio: Which Hypotheses Remain? A Systematic Review. *Front. Psychol.* 10:1221. doi: 10.3389/fpsyg.2019.01221

Over the last 25 years, a large amount of research has been dedicated to identifying men's preferences for women's physical features, and the evolutionary benefits associated with such preferences. Today, this area of research generates substantial controversy and criticism. I argue that part of the crisis is due to inaccuracies in the evolutionary hypotheses used in the field. For this review, I focus on the extensive literature regarding men's adaptive preferences for women's waist-to-hip ratio (WHR), which has become a classic example of the just-so storytelling contributing to the general mistrust toward evolutionary explanations of human behavior. The issues in this literature originate in the vagueness and incompleteness of the theorizing of the evolutionary mechanisms leading to mate preferences. Authors seem to have rushed into testing and debating the effects of WHR on women's attractiveness under various conditions and using different stimuli, without first establishing (a) clear definitions of the central evolution concepts (e.g., female mate value is often reduced to an imprecise concept of "health-and-fertility"), and (b) a complete overview of the distinct evolutionary paths potentially at work (e.g., focusing on fecundability while omitting descendants' quality). Unsound theoretical foundations will lead to imprecise predictions which cannot properly be tested, thus ultimately resulting in the premature rejection of an evolutionary explanation to human mate preferences. This paper provides the first comprehensive review of the existing hypotheses on why men's preferences for a certain WHR in women might be adaptive, as well as an analysis of the theoretical credibility of these hypotheses. By dissecting the evolutionary reasoning behind each hypothesis, I show which hypotheses are plausible and which are unfit to account for men's preferences for female WHR. Moreover, the most cited hypotheses (e.g., WHR as a cue of health or fecundity) are found to not necessarily be the ones with the strongest theoretical support, and some promising hypotheses (e.g., WHR as a cue of parity or current pregnancy) have seemingly been mostly overlooked. Finally, I suggest some directions for future studies on human mate choice, to move this evolutionary psychology literature toward a stronger theoretical foundation.

Keywords: mate choice, attractiveness, evolutionary hypotheses, WHR, mate value, reproductive success, fertility

INTRODUCTION

The ratio between the waist and the hips circumferences (Waist-to-Hip Ratio, or WHR) is a physical characteristic often used as an example to show that evolution shaped human mate preferences. It is also an example of just-so storytelling in evolutionary psychology. In 1993, Devendra Singh suggested that WHR represents a strong predictor of women's physical attractiveness (Singh, 1993a). He also argued that men's preference for a mate with a low WHR is adaptive, because a low WHR reflects a woman's high mate value. But what exactly is this "mate value"? During the past 25 years, the evolutionary literature on WHR and women's attractiveness has flourished, but the definition of this "mate value" is rarely expressed. In evolutionary biology, mate value is attached to the concept of reproductive success: a woman with a high mate value will increase the reproductive success of her mate(s). An increase in reproductive success is characterized by an increased number of descendants in next generations and can be achieved in various ways. First, survival until reproduction is indispensable. Second, the number of children born during an individual's lifespan is also crucial. But the survival and the quality of these children will directly impact their own reproductive success, and hence the number of grandchildren in the next generation, thus ultimately influencing the reproductive success of the grandparents. In short, a woman has higher value as a potential mate if she increases the number and quality of descendants a man will have (including the ones he has with other women). The question then is which of these components of reproductive success are actually linked to a mate's WHR? To answer this, I assemble the numerous hypotheses exposed since the idea of the WHR as an indicator of women's mate value was first suggested in 1993. These hypotheses are examined to determine which of the characteristics linked to WHR are most likely, in theory, to be translated into an increase in the reproductive success of the woman's mate.

The objective of this review is 2-fold. The first goal is to gather and pool all the existing evolutionary hypotheses regarding men's preferences for a certain (low, high or average) WHR. There are many reviews about men's preferences for women's WHR, but this is the first exhaustive review of the hypotheses mentioned in these studies. The second purpose of this paper is an in-depth theoretical examination of these hypotheses, which are often only briefly justified and, in some cases, have never been properly developed.

Most of the debate around WHR and attractiveness has centered on two other questions: "Is the preference for a low WHR universal?" and "Is WHR the best predictor of the attractiveness of women's bodies?" I will not address these two questions extensively here (it is beyond the scope of this paper), but a brief commentary seems necessary at this point. A preference for a relatively low WHR (i.e., low relatively to men's WHR, or low relatively to the average female WHR) has been observed in a large number of studies, including a wide range of populations and methods. With that in mind, results show that there is some variation in what is the exact value of the ideal WHR [reviewed in Brooks et al. (2015) and Cashdan (2008)]. The second debate concerns WHR as

the "best" predictor for attractiveness. Authors have debated whether WHR or BMI is the best predictor of attractiveness and mate value (Tassinari and Hansen, 1998; Tovée et al., 1999; Furnham et al., 2005; Cornelissen et al., 2009a,b). As could be expected, the results vary according to the population and stimuli used. Other measurements have also been proposed to replace WHR (for example, hip or waist size alone, abdominal depth or waist/stature ratio: Brooks et al., 2010, 2015; Lassek and Gaulin, 2016). The objective of this paper is not to decide if WHR is the best measure of physical attractiveness or if the ideal WHR is universal or not. For our purposes, it is sufficient to note that the effect of WHR on attractiveness is widespread (even if the value of the preferred WHR varies), and large enough to warrant questions about its possible adaptive basis.

MATERIALS AND METHODS

The dataset encompasses any articles and book chapters addressing men's preferences for women's WHR, based on an evolutionary approach (see the **Supplementary Material** for the details). The final dataset consists of 104 papers from 58 different first authors, including 13 review papers and chapters from 1993 to 2017. All the hypotheses concerning men's adaptive preferences toward women's WHRs, waist size or hip size are collected (see **Figure 1** and **Supplementary Table S1**).

See the **Supplementary Material** for the details about the methodology used in the collection and the selection of hypotheses.

RESULTS

In the following sections, I review each hypothesis found in the literature to see if it could, in theory, support an adaptive role of the preference for a low WHR. For a hypothesis to be plausible, two steps are required: 1) Correlation with WHR: first, WHR needs to be correlated with the biological trait of interest (Is WHR associated with the nominated characteristic in the population?). This correlation needs to be strong enough, such that a detectable variation of WHR attractiveness translates into a significant variation of the hypothesized trait; 2) Effect on the man's reproductive success: second, the nominated characteristic should be associated with a potential increase of reproductive success (meaning more descendants, and higher quality descendants) for the individual who chooses a mate carrying this characteristic. A dispensable third step can be added: 3) Perception of the characteristic using WHR: do people use the WHR to assess the nominated characteristic (Are people conscious of the link between WHR and the characteristic)? Importantly, this third step is not mandatory for the hypothesis to be valid, as people do not need to be conscious of the biological consequences of their preferences for them to have an effect. In other words, a preference for a trait can perfectly evolve when individuals do not suspect that this trait is a cue of something else. For example, a preference for sweet taste evolved in our ancestors without them knowing that it was a cue of a source of readily available energy. However, this third step can

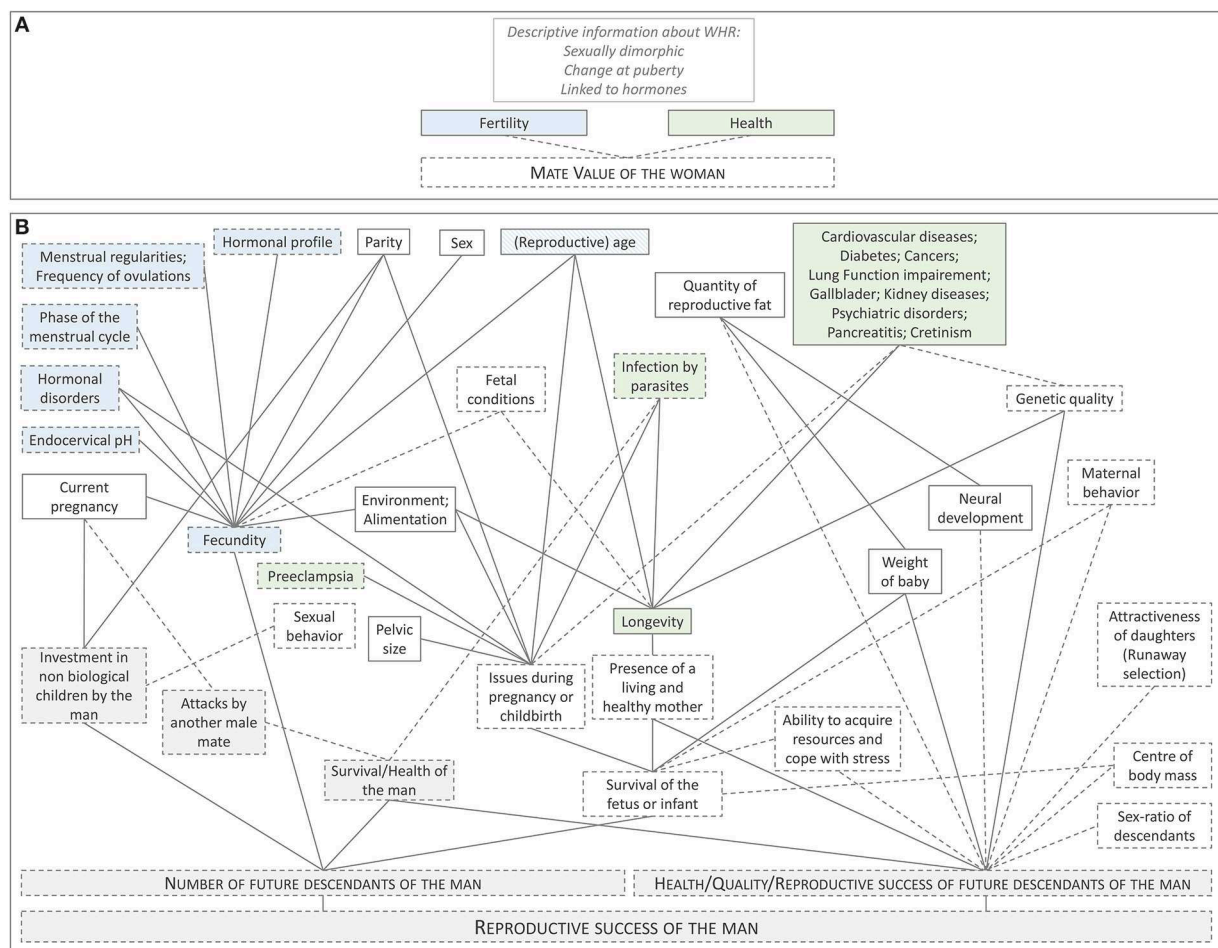


FIGURE 1 | Two theoretical frameworks explaining men's adaptive preferences for women's WHR. **(A)** Example of a vague theoretical explanation often found in the literature. WHR is assumed to be a cue of women's health and fertility, which supposedly translate into women's "mate value." Some descriptive information of WHR is sometimes added to the theory (e.g., WHR is sexually dimorphic), but without any explicit link to mate value. **(B)** A more complex but more accurate theoretical explanation, including all the different hypotheses found in the literature. Each box represents a characteristic linked to women's WHR. The diagram illustrates their potential links to the reproductive success of the man (the woman's mate). The boxes with a gray background directly concern the man. All the other boxes relate to the woman's characteristics. Characteristics related to women's fertility (as usually defined in this literature) are represented in blue. In green: characteristics related to women's health. The lines connecting the boxes represent correlations, without implying any causality. Dotted lines indicate that empirical evidences for the correlation are scarce. Dotted frames indicate that evidences linking the characteristic to women's WHR are scarce. The links between each characteristic are not represented on the diagram. For example, parity is obviously correlated with a woman's age, but this correlation is not illustrated here (each characteristic is supposed to be correlated to WHR, controlling for other characteristics).

represent additional support in favor of the hypothesis and help us understand the mechanisms behind mate preferences.

Cue of Biological Sex

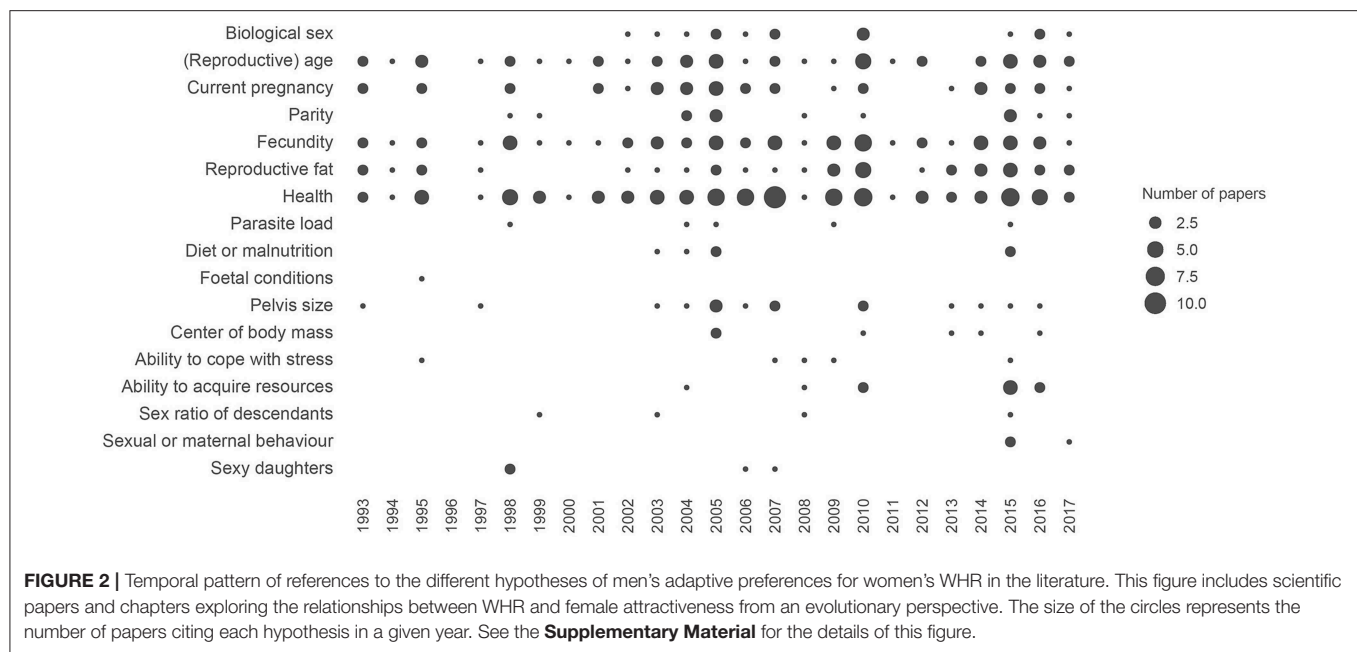
According to this hypothesis, WHR would be a way to detect the biological sex of a potential mate. The first mention of this straightforward hypothesis appears relatively late, almost 10 years after the first paper about adaptive preferences for WHR (Tovée et al., 2002, see **Figure 2**), and is present in only 14% of the papers (see **Supplementary Material**).

Correlation With WHR

WHR is sexually dimorphic in the human species. A significant difference between men's and women's WHR has been found

in all the populations where it has been investigated (Leibel et al., 1989; Björntorp, 1991, 1993; Marti et al., 1991; Beall and Goldstein, 1992; Ley et al., 1992). The size of the difference between sexes varies between populations, but no culture has been found where men have a lower average WHR than women. Thus, WHR is a reliable cue of biological sex.

However, there are many other traits which are sexually dimorphic in humans and which can be used to assess biological sex: height, shoulder-to-hip ratio, hair, facial traits, breast, genital, voice, and so on (Wenzlaff et al., 2018). Consequently, people can identify biological sex without using WHR. And because WHR is not indispensable to assess sex, the selective advantage of a preference for a low WHR as a way to assess biological sex is



reduced (Iwasa and Pomiankowski, 1994; Bro-Jørgensen, 2010). On the other hand, the “redundant signaling” hypothesis (or “back-up signal” hypothesis) claims that multiple cues conveying similar information (the biological sex in the present case) compensate for errors during information coding (Møller and Pomiankowski, 1993; Johnstone, 1996). In other words, multiple cues serve as a back-up signal that ensures a low rate of mate choice errors (Bro-Jørgensen, 2010; Abend et al., 2015). Moreover, when the ability to detect different cues varies with environmental conditions or distances, individuals may pay attention to different cues under different conditions (Candolin, 2003). More experiments are required to measure by how much and when WHR does improve the detection of biological sex, in addition to other sexually dimorphic features.

Effect on the Man's Reproductive Success

Obviously, it is necessary to copulate with the opposite sex to increase the number of descendants. The problem in this case is not the effect of the characteristic (biological sex) on the individual's reproductive success, but the non-uniqueness of the cue. Thus, unless we discover evidence supporting the “redundant signaling” hypothesis, the sexual dimorphism of WHR may have contributed to the selection of men's preference for a low WHR but was probably not the only selective force involved.

Perception of the Characteristic Using WHR

People use WHR to assess individuals' biological sex, and a lower WHR is strongly associated with figures being perceived as female (Johnson and Tassinari, 2005; Johnson et al., 2010; Saunders et al., 2010; Pazhoohi and Liddle, 2012). People are able to detect sex using WHR when other cues of sex are unavailable (Pazhoohi and Liddle, 2012) but, as stated earlier, we need to measure the accuracy of this detection with and without the use of WHR. This

would give us an indication of the strength of the selection on the use of WHR as a cue of biological sex.

Cue of (Reproductive) Age

This hypothesis, already referred to in the first paper on the topic (Singh, 1993a, see **Figure 2**), is found in 43% of the papers. According to this hypothesis, WHR would be an indicator of chronological or reproductive age.

Correlation With WHR

WHR is high in early childhood (and similar between boys and girls) and drops around the onset of puberty for women. Then, women's WHR increases from the peak of fertility (in the 20's). This general age pattern is observed in many countries, including non-western and non-industrial populations (Rimm et al., 1988; Leibel et al., 1989; Seidell et al., 1990; Beall and Goldstein, 1992; Ley et al., 1992; Björntorp, 1993; Casey et al., 1994; Sugiyama, 2004; Bohler et al., 2010; Brooks et al., 2010; Bacopoulou et al., 2015; Butovskaya et al., 2017). Finally, WHR might also be a cue of menopause (end of reproductive age, independently of chronological age; Kirschner and Samojlik, 1991; Björkelund et al., 1996; Tchernof and Poehlman, 1998), but several other studies find no effect of menopausal status on WHR (Lanska et al., 1985; Tonkelaar et al., 1989; Seidell et al., 1990; Troisi et al., 1995; Tchernof and Poehlman, 1998; Sugiyama, 2004).

To sum up, WHR is a reliable cue of the start of women's reproductive capacity (menarche and puberty). WHR is also a reliable indicator of women's age after puberty, and maybe of menopause. However, as with biological sex (although to a lesser extent), individuals can rely on other cues to assess age (the face, for example, and menarche is also linked to an increase in breast size). The redundancy of the cue decreases the selective advantage of the preference for a low WHR as cue of age (Iwasa and Pomiankowski, 1994; Bro-Jørgensen, 2010). Nevertheless, according to the “redundant signaling” hypothesis, using several

cues simultaneously, including WHR, may increase the precision of the age estimation (Johnstone, 1996; Bro-Jørgensen, 2010; Abend et al., 2015). Alternatively, the use of redundant cues may reduce time and energy spent inspecting mates, make mate assessment possible under different conditions (Rowe, 1999; Candolin, 2003), or make it more difficult for women to cheat about their actual age (Candolin, 2003).

Effect on the Man's Reproductive Success

Fecundability (the ability to become pregnant) is age-dependent for women. Fecundability is null before menarche, increases from puberty, peaks in mid-twenties on average, and then decreases until menopause, the end of the reproductive window (Menken and Larsen, 1986; Weinstein et al., 1990; Dunson et al., 2002; Wallace and Kelsey, 2010). In addition, the risks of complications during pregnancy and childbirth are also related to age (more issues when very young, and after 30, even when controlling for parity; Naeye, 1983; Fretts et al., 1995; Amarín and Akasheh, 2001). Thus, the choice of a mate around the peak of fertility (for short-term relationships), or before (for long-term relationships), will increase the number of potential descendants a man can sire with this mate, and thus his reproductive success.

However, because WHR is not the only cue of women's age, the correlation between age and WHR may have contributed to the selection and/or maintenance of men's preferences for a low WHR, but it is unlikely to be sufficient on its own, unless the use of WHR in addition to other redundant cues increases the precision of the age estimation in a way which would confer a selective advantage to the men making this estimation. More investigation is necessary to explore the role of the back-up signal hypothesis in this case.

Perception of the Characteristic Using WHR

People seem to use WHR to assess women's age, and a low WHR tends to be associated with perceptions of youthfulness (Singh, 1993b, 1994, 1995; Singh and Luis, 1995; Furnham et al., 2004, 2005; Sugiyama, 2004; Andrews et al., 2017). However, many other results on perceived age are inconclusive (Singh, 1993a,b; Henss, 1995, 2000; Singh and Young, 1995; Furnham et al., 1997, 2002, 2005; Sorokowski et al., 2014; Wang et al., 2015). I suggest this is due to two reasons: first, WHR does not have a linear relationship with age (it is U shaped), and secondly, people simultaneously use other cues to infer age. As a consequence, depending on other cues depicted on the stimuli (face, breasts or hair, for example), results can reveal a negative, positive, or null relationship between perceived age and WHR. Further studies investigating the interaction effect between WHR and other physical cues on perceived age are needed. Moreover, the effect of WHR on perceived youthfulness should be explored in different populations, as most of the studies have been conducted in WEIRD countries (but see Furnham et al., 2002; Sugiyama, 2004; Sorokowski et al., 2014 for notable exceptions).

Cue of Current Pregnancy

According to this hypothesis, a woman's WHR could be used to detect if she is currently pregnant or not. This hypothesis is found in 31% of the papers.

Correlation With WHR

Women experience a drastic increase in WHR during pregnancy, which is mainly due to an increase in waist circumference. A decrease in hip circumference may also happen, as fat from this region is mobilized during late pregnancy to meet the needs of the growing fetus (Rebuffé-Scrive et al., 1985; Lassek and Gaulin, 2006). Moreover, unlike for sex and age, WHR is the unique reliable visual cue of pregnancy. The slope of the decreasing attractiveness of WHR through pregnancy for different populations remains to be specified. The earlier in pregnancy the WHR starts to be significantly less attractive, the more plausible this hypothesis will be, as men would be able to use this cue longer/more often.

Effect on the Man's Reproductive Success

Because women are infertile while pregnant, pregnancy is directly linked to current fecundity. As such, choosing a pregnant woman as a short-time mate will not enhance a man's reproductive success. However, being pregnant is a transient stage. For a young woman, pregnancy is positively associated with her future expected reproductive success (it is a sign that she is able to become pregnant and to carry a child). However, for older women, being currently pregnant is negatively correlated with the expected number of additional children. Some authors argue that the relationship between current pregnancy and future fertility depends on the fertility rate of the population: if the total fertility rate of the population is low (e.g., two children) it would be costly to be attracted to a woman who is already pregnant, because there is a high risk that she may conceive only one more child. In traditional societies, where total fertility rates sometimes exceed 6, if a woman is pregnant, she may nevertheless conceive at least a few more children (Marlowe and Wetsman, 2001). To sum-up, choosing a mate who is already pregnant will, most of the time, decrease the number of potential descendants of an individual, but the size of the effect depends on the woman's age, the total fertility rate in the population and the type of relationship (short or long term).

Moreover, choosing a long-term mate who is pregnant with the child of another man entails additional evolutionary costs, because investing in a non-biological child decreases the amount of investment an individual can invest in his own descendants. Lastly, Singh suggests that choosing a mate carrying the baby of another man could increase the risks of violence from the jealous current mate (which could impact the survival or future reproductive success of the individual suffering from the attack).

Altogether, mating with a woman with a high WHR because it indicates current pregnancy will have, on average, a negative effect on an individual's reproductive success. Added to the fact that WHR is a reliable and distinctive cue of current pregnancy, this gives solid theoretical support in favor of the present hypothesis.

Perception of the Characteristic Using WHR

To my knowledge, only two studies have investigated the role of WHR on the perception of current pregnancy (Furnham et al., 2001; Schützwohl, 2006). The results confirm that a low WHR is associated with a lower perceived probability of current

pregnancy. The perception of pregnancy using WHR seems obvious for the last stages of pregnancy, based on profile views or in 3D. However, and even if the conscious awareness of a pregnancy is not mandatory for men's preference to evolve, it would be interesting to explore when exactly people start to detect pregnancy, using 2D images including frontal, back and profile views of women at different pregnancy stages.

Cue of Parity (Number of Previous Pregnancies)

This hypothesis, first mentioned in 1998 (Yu and Shepard, 1998, see Figure 2), is present in 11% of the papers in this literature. It stipulates that WHR is a way to estimate the number of children (or number of pregnancies) that a woman has previously had in her life.

Correlation With WHR

There is evidence that WHR increases with the number of previous pregnancies (independently of age and BMI), due to an increase of waist circumference and/or a relative decrease in hip circumference (Kaye et al., 1990; Smith et al., 1994; Troisi et al., 1995; Bjorkelund et al., 1996; Rodrigues and Costa, 2001; Lassek and Gaulin, 2006; Wells et al., 2010, 2011; Butovskaya et al., 2017). This change in body shape (sometimes referred to as (covert) maternal depletion) is due to the mobilization of fat from the lower parts of the body to meet the needs of the developing child (as well as looser abdominal muscles). This may be interpreted as a life history strategy for allocating energy between competing gluteofemoral fat depots for reproduction, and central fat depots for maintenance and survival (Cashdan, 2008; Wells et al., 2010, 2011). This phenomenon has been observed in various countries: Brazil (Rodrigues and Costa, 2001), Sweden (Bjorkelund et al., 1996), Thailand (Wells et al., 2011), UK (Wells et al., 2010), USA (Kaye et al., 1990; Troisi et al., 1995; Lassek and Gaulin, 2006), and non-industrial societies including tribes from Sub-Saharan Africa, Western Siberia, South America and South Asia (Butovskaya et al., 2017). However, a few other studies find that parity has a negligible or null effect on WHR (Lanska et al., 1985; Tonkelaar et al., 1989; Seidell et al., 1990; Nenko and Jasienska, 2009), but these null results can be explained by the higher average age of the women sampled in those studies. Indeed, the parity effect seems to dissipate over time (Wells et al., 2010). Note that this does not affect the plausibility of the present hypothesis, as the effect of parity on WHR should be visible at the time of mate choice (relatively young).

Effect on the Man's Reproductive Success

Women's limited reproductive potential and resources mean that, even controlling for age, each child already born reduces the future number of children a man can sire with the woman if he mates with her long-term (Symons, 1981; Sugiyama, 2005). Parity status influences the survival and quality of future descendants. For example, both high parity and nulliparity are associated with increased risks during childbirth and lower birthweights (Kiely et al., 1986; Fretts et al., 1995; Hinkle et al., 2014; Merklinger-Gruchala et al., 2015, 2017), and IQ is negatively correlated with birth order (Downey, 2001). A recent

pregnancy also increases the probability of current infertility because of lactational amenorrhea. Finally, as with current pregnancy, higher parity increases the costs linked to investment in genetically unrelated children.

In conclusion, even when the risks associated with first births are taken into account, choosing a mate with a low parity should have an overall positive impact on individuals' reproductive success (especially for long term relationships), and WHR as a cue of parity is likely to play a significant role in the selection of men's preferences for a low WHR.

Perception of the Characteristic Using WHR

To my knowledge, only one study investigates the effect of WHR on perceived parity, with the results validating that women with a higher WHR are perceived as having a higher number of children (Andrews et al., 2017). This study needs replications in populations other than undergraduate students from the USA, but the results suggest that people are using WHR as a cue of parity.

Cue of Fecundity

One of the most cited argument for an adaptive preference for a low WHR is WHR as a cue of fecundity (cited in 54% of the papers). Healthy women of similar age and reproductive history vary in their ability to become pregnant and achieve a live birth, and WHR would be an indicator of this ability.

Correlation With WHR

The most direct evidences in favor of this hypothesis comes from a few clinical studies showing that women with a lower WHR have a higher probability of conception in the case of *in vitro* fertilization and artificial insemination (Zaadstra et al., 1993; Wass et al., 1997). But more recent studies find no relationships between women's WHR and their likelihood of conceiving after induction of ovulation (Imani et al., 2002; Eijkemans et al., 2003). These studies are informative because they are directly linked to fecundity, but women seeking medical assistance to conceive do not represent the ideal population to investigate factors of natural fecundity.

A few studies find that high WHRs are correlated with a later age at first live birth (Kaye et al., 1990) or longer time-to-pregnancy (Wise et al., 2013; McKinnon et al., 2016; but see Wise et al., 2010).

An indirect way to detect the link between fecundity and WHR is to look at the menstrual cycles or at the physiological factors linked to both WHR and fecundity. A few studies indicate that WHR is linked to menstrual abnormalities (Hartz et al., 1984; Moran et al., 1999) and to hormonal levels linked to fecundity (Björntorp, 1991; Jasienska et al., 2004). Similarly, one study finds that women with low WHRs have lower endocervical pH (Jenkins et al., 1995), which helps sperm penetration (Zavos and Cohen, 1980). However, these results seems not to hold for non-obese young women (see Lassek and Gaulin, 2018b for a richer discussion on this topic).

Finally, one study finds that WHR decreases around ovulation (Kirchengast and Gartner, 2002), suggesting that WHR might also reveal whether a woman is at peak cycle fertility. However,

these results should be interpreted with caution, as others fail to replicate this effect (Bleske-Rechek et al., 2011).

To conclude, there are some indirect lines of evidence that WHR could be linked to fecundity, but this effect is mostly found when high WHR is associated with other factors (as obesity or older age) and might thus be negligible in populations of young and non-obese women (Lassek and Gaulin, 2018b). Moreover, these studies almost exclusively focus on WEIRD populations, limiting even more the generalization of these results.

Effect on the Man's Reproductive Success

Choosing highly fecund mates will increase the reproductive success of a man both for long-term and short-term relationships. In the case of a short-term relationship, it will simply increase the probability of a pregnancy. In the case of a long-term relationship, it will increase the number of potential descendants by reducing both interbirth intervals and the period before the first child (thus increasing the reproductive window).

However, in light of the lack of evidence of a link between WHR and young and non-obese women's fecundity, this hypothesis does not benefit from strong empirical support.

Perception of the Characteristic Using WHR

A few studies find that a low WHR is associated with higher perceived fecundity (Singh, 1993b; Furnham et al., 2004; Sugiyama, 2004), but the results are unclear for the vast majority of the cases (Singh, 1993b, 1994, 1995; Singh and Luis, 1995; Furnham et al., 1997, 2001, 2003, 2004, 2005, 2006; Tassinari and Hansen, 1998). I suggest that this lack of clarity is mainly due to the ambiguity of the questions asked to the participants. The main issue is the absence of any indication about the time frame. For example, high parity (linked to a high WHR), is positively associated with past fecundity, but negatively associated with future fecundity [see section Cue of Parity (Number of Previous Pregnancies)]. Thus, in the absence of additional information, it is impossible to know if the participants are rating past, current or future fecundity. The answer probably depends on other cues provided in the survey, or vary from one participant to another, which could explain the inconclusive results. Future tests of perceived fecundity should include the notion of time.

Cue of Quantity and Availability of "Reproductive Fat"

The idea that fat located around women's hips is qualitatively different from fat found in other body regions, and is used specifically for reproductive functions, exists in the literature since 1993 (Singh, 1993b, see **Figure 2**). This hypothesis has been progressively enriched, stating that a mother's WHR is linked to the development of her fetus and infant. It is present in 34% of the papers.

Correlation With WHR

WHR is, by construction, positively correlated with the quantity of fat situated at the waist level (abdominal fat), and negatively correlated with fat quantity located around the hip

(gluteofemoral fat). There is evidence that gluteofemoral fat in women is specific to reproduction: the storing of gluteofemoral fat is high (compared to males and to other body parts) during human female development (Fredriks et al., 2005). Moreover, even with restricted food intake, gluteofemoral fat is metabolically protected from use until late pregnancy and lactation, when it is selectively mobilized (Rebuffé-Scrive et al., 1985; Lassek and Gaulin, 2008). The hypothesis derived from these observations is that the quantity of gluteofemoral fat would have an effect on the development of the fetus during pregnancy and of the infant through lactation.

This reproductive fat appears to be of particular importance for brain development, as gluteofemoral fat is the main source of long-chain polyunsaturated fatty acids that are critical for fetal and infant neural development. Additionally, it seems that abdominal fat inhibits the availability of these neurodevelopment resources (abdominal fat decreases the amount of the enzyme Δ -5 desaturase, which is rate limiting for the synthesis of long-chain polyunsaturated fatty acids; Lassek and Gaulin, 2008). Consequently, WHR is an indicator of the quantity and availability of the fatty acids needed for fetal and infant brain development. In favor of this hypothesis, a study shows that women with lower WHRs and their children have significantly higher cognitive test scores (Lassek and Gaulin, 2008).

Moreover, one study finds that a low WHR correlates with higher birth weight (Pawłowski and Dunbar, 2005), but other studies found the opposite (Brown et al., 1996; Salem et al., 2012).

To conclude, a woman's WHR seems to be a promising indicator of future fetus and infant neural development (although further data from different countries are needed), and additional evidence is required to confirm the link between pre-pregnancy WHR and fetal growth.

Effect on the Man's Reproductive Success

Mating with a woman able to provide enough resources during the development of the fetus and infant increases the survival and quality of the descendants. Offspring with higher cognitive abilities are likely to have a better rate of survival and reproductive success than individuals who suffer from worse conditions during their brain development.

A low birthweight is associated with higher infant mortality (Chase, 1969; Behrman et al., 1982; McCormick, 1985) and negative outcomes later in life (Hackman et al., 1983; Baker et al., 2008). However, a low birthweight is also associated with variables which may have no effect on the father's reproductive success (e.g., because occurring late in life), and could even have a positive effect in some environments (Bateson et al., 2004), as a low birthweight seems to be associated with a faster life history strategy (Nettle, 2010).

In conclusion, choosing a mate with a lower WHR if it is linked to higher resources for fetal and infant brain development (and maybe general growth), will have a generally positive impact on a man's reproductive success. However, the size of this effect according to the environmental conditions should be explored. For example, how does this trait impact the number of descendants in the next generation when conditions are more favorable to faster life history strategies?

Perception of the Characteristic Using WHR

To my knowledge, only one study explores the effect of WHR on the perceived quality of the descendants (Andrews et al., 2017). Andrews et al. (2017) ask participants to rate female bodies for the following questions: “If this woman were to have a child, it would be healthy;” “If this woman were to have a child, it would make friends easily;” “If this woman were to have a child, it would be popular.” They find a negative relationship between WHR and projected offspring quality, supporting the idea that women with low WHRs are expected to have higher quality children than women with high WHRs (but, as often with this type of questions, it is difficult to tell if we are measuring something else than a halo effect).

Cue of Health

One of the most cited hypotheses stipulates that a low WHR is an indicator of women’s good health (hypothesis present in 87% of the papers). The health conditions which are referred to in the literature on WHR and attractiveness are: cardiovascular diseases, hypertension, strokes, myocardial infarction, diabetes, gallbladder disease, kidney diseases, pancreatitis, lung function impairment, cretinism, psychiatric disorder, various cancers and preeclampsia.

Correlation With WHR

A high WHR is correlated with many health issues. This claim is supported by abundant evidence (for reviews see Björntorp, 1987a,b, 1993; Manolopoulos et al., 2010). However, these findings are based on relatively old women or men (often 60 years old or more, almost never before 30), mostly suffering from some degree of obesity, raising the possibility that this relationship is not present for evolutionary relevant reproductive-age populations (Lassek and Gaulin, 2018a).

Effect on the Man’s Reproductive Success

The consequences for reproductive success of mating with a woman with a low WHR because it is a cue of her health are not straightforward. First, the cited health conditions are not contagious, thus the survival of the woman’s mate cannot be directly affected. Secondly, most of the chronic diseases associated with WHR are recent, from an evolutionary point of view, and they are associated with present-day environments, lifestyle and alimentation (Eaton and Eaton, 1997; Groop, 2000). Third, even if we assume that these health issues were common in our evolutionary past, most of them appear late in life, after the end of women’s reproductive life. Thus, most of the health issues linked to high WHRs are unlikely to affect the number of descendants of a woman’s mate (Lassek and Gaulin, 2018a).

A few exceptions in the list of WHR-related health issues can be made, however. First, a high WHR early in pregnancy seems to be correlated to higher risks of preeclampsia (a condition which can be fatal to both the fetus and mother; Yamamoto et al., 2001; Taebi et al., 2015). However, evidence is needed to see if preeclampsia is predicted by WHR *before* pregnancy (when mate choice occurs). One paper indicates that a high WHR can be an indicator of cretinism (a syndrome often linked to infertility; Streeter and McBurney, 2003). However, WHR

is probably not a very good cue to detect cretinism, as this health condition generates other physical modifications, more easily noticeable than WHR (Chen and Hetzel, 2010). Another exception is the polycystic ovarian syndrome. This condition can affect the fertility of young women, but only when the syndrome is associated with obesity (Pall et al., 2006; Pasquali et al., 2006). And again, the prevalence of this condition in our evolutionary past is unclear. Lastly, the term “health” can include malnutrition and parasites (although it is almost never referred to in the literature), which can affect fertility at any age and are not restricted to our contemporary societies. These two last characteristics are discussed in the next sections of this paper (Cue of Parasite Load & Cue of Diet).

Health later in life could influence the survival and quality of descendants in another way, through maternal investment: long-term health and longevity increase the probability of having a living and healthy mother able to provide care for children and grandchildren (Sear et al., 2000). Thus, theoretically, WHR as a cue of health could have played a role in the selection of preferences for a low WHR. However, this hypothesis holds only if WHR at a younger age (at the time of mate choice) is a reliable predictor of health later in life, excluding diseases which are evolutionary novel. Longitudinal studies in non-WEIRD populations are needed to explore this possibility.

Alternatively, good health at old age could be related to genetic quality. Descendants from individuals with higher longevity could have a better health, even at younger ages. In this case, men’s preferences for a low WHR as a cue to health could evolve through indirect selection. Cross-generational studies are needed to test this good genes hypothesis.

To conclude, in the light of the present evidence, the “WHR as a cue of health” hypothesis is unlikely to be at the evolutionary origins of preferences for a low WHR in young women. However, this hypothesis could receive new theoretical support through the maternal and grandmaternal investment or the genetic quality hypotheses, but only if some of the above predictions (links between women’s WHR at young age and health at old age, or health of the descendants, excluding evolutionary novel diseases) are supported by evidence.

Perception of the Characteristic Using WHR

Participants are asked to rate the health of the stimuli in many studies (Singh, 1993a,b, 1994, 1995; Singh and Luis, 1995; Singh and Young, 1995; Furnham et al., 1997, 1998, 2001, 2002, 2003, 2004, 2005, 2006; Yu and Shepard, 1998; Wetsman and Marlowe, 1999; Henss, 2000; Marlowe and Wetsman, 2001; Sugiyama, 2004; Marlowe et al., 2005; Schützwohl, 2006; Tovée et al., 2007; Swami et al., 2009; Sorokowski et al., 2014). In general, a low WHR is associated with better perceived health. Interestingly, however, a few studies investigating non-WEIRD populations find a null or positive effect of WHR on perceived health (Yu and Shepard, 1998; Wetsman and Marlowe, 1999; Tovée et al., 2007; Sorokowski et al., 2014). This supports the idea that the association between high WHR and poor health might be valid in contemporary western countries only. Even if, as explained earlier, the perception of health using WHR is not a mandatory

step to validate the hypothesis, more research (with different stimuli and questions) is needed to clarify this point.

It would also be interesting to see if young women's WHR is linked to their perceived future health and longevity. One could also explore if individuals have any idea of the kind of diseases associated with WHR.

To my knowledge, only one study explores the effect of WHR on the perceived quality of the descendants (Andrews et al., 2017, see section Cue of Quantity and Availability of "Reproductive Fat" above). They find a negative relationship between women's WHR and the projected offspring quality, in accordance with the hypothesis of WHR as a cue of genetic quality.

Cue of Parasite Load

The idea that WHR could be a sign of infection by parasites is not recent (e.g., Furnham et al., 1998, see **Figure 2**) but is quite rare in the literature (in 5% of the papers).

Correlation With WHR

Some parasites, including intestinal worms, can increase waist size through oedema while causing weight loss, which will result in a higher WHR (Cross, 1992; Kucik et al., 2004).

Effect on the Man's Reproductive Success

Parasite load can affect survival and fertility. Moreover, most parasites are contagious, and mating with a woman carrying parasites increases the probability of being infected. As such, WHR as a cue of parasite load can have an effect on a man's health and survival, as well as an effect on the number, survival and quality of descendants he can sire with the infected woman. This effect remains to be quantified and will certainly vary according to the frequencies and types of parasites present in the environment.

WHR as a cue of parasite load is an interesting hypothesis, but it has been largely overlooked and evidence is by consequence lacking.

Perception of the Characteristic Using WHR

There is no specific research on the perception of parasite load based on WHRs. However, many studies explore the effect of WHR on perceived general health (see section Cue of Health).

Cue of Diet or Malnutrition

The hypothesis that WHR could be a cue of women's diet or malnutrition is found in 5% of the papers.

Correlation With WHR

One paper mentions that a high WHR could be a sign of Kwashiorkor, a form of malnutrition (Streeter and McBurney, 2003). Indeed, WHR can increase in some cases of malnutrition because of the presence of an oedema enlarging waist size (Golden, 1982; Waterlow, 1984).

A diet rich in fibrous food can also increase waist size and thus WHR. For example, Marlowe states that Hadza women may have a high WHR because "a larger gut is required to hold the amount of bulky, fibrous tubers in the Hadza diet" (Marlowe et al., 2005).

Effect on the Man's Reproductive Success

Malnutrition increases the morbidity and mortality of a woman and her children, and might also decrease her fecundity (Mosley, 1977; Osteria, 1982; Hernández-Julián et al., 2014). Choosing a mate suffering from malnutrition will thus decrease one's reproductive success. The prevalence of malnutrition involving a high WHR during our evolutionary past should be explored, to establish if it could have represented an evolutionary force for the preferences toward low WHRs.

Concerning diet, it is not clear if a large waist reveals a good ability to digest fibrous food or a poor ability to assimilate this kind of food. If the latter is true, a higher WHR will be associated with less resources available for pregnancy and lactation, leading to lower survival and quality of descendants. The opposite will be true if a large waist is associated with a better ability to digest fibrous food.

The hypotheses of WHR as a cue of malnutrition or diet (or ability to digest some type of food) have been mainly ignored, and evidence is thus missing.

Perception of the Characteristic Using WHR

There is no specific research on the perception of diet or malnutrition based on WHR.

Cue of Fetal Conditions

This hypothesis is mentioned only once in the literature (Singh, 1995). It stipulates that the WHR of an adult woman could be an indication of her developmental conditions before her birth.

Correlation With WHR

A negative link between adult WHR and birth weight, or placental weight to birth weight ratio (an indicator of retarded fetal growth), has been found, but this study is only composed of men over 50 years old (Law et al., 1992). To my knowledge, there is no empirical evidence showing that young women's WHR is a reliable cue of their fetal development.

Effect on the Man's Reproductive Success

A low birthweight is associated with higher infant mortality (Chase, 1969; Behrman et al., 1982), but this cannot affect a mate's reproductive success, as the mating occurs after the woman's survival to infancy. But a low birthweight also has some negative outcomes later in life (Bateson et al., 2004), for women's fertility (Hackman et al., 1983) and longevity (Baker et al., 2008, which decreases the likelihood of having a living and healthy mother caring for her mate's descendants, see section Cue of Health).

On the other hand, as explained in section Cue of Quantity and Availability of "Reproductive Fat," a low birthweight is also associated with some advantages in harsh environments (Bateson et al., 2004), as well as a relatively early sexual maturation and reproduction (Nettle, 2010), which might increase the number of descendants for the potential mate.

To conclude, WHR as a cue of a woman's fetal condition could have, in theory, a negative, positive or null effect on her mate's reproductive success. Combined with the fact that the link between WHR and fetal conditions has been shown

for older men only, this hypothesis lacks both empirical and theoretical support.

Perception of the Characteristic Using WHR

There is no test of the effect of WHR on perceived fetal conditions.

Cue of Pelvis Size

This hypothesis, found in 16% of the papers in this literature, is already reported in one of the first papers from Singh (1993b), see **Figure 2**, and states that WHR is a cue of the size (or shape) of women's pelvis.

Correlation With WHR

WHR is, by definition, linked to hip size, which is indicative of underlying pelvic skeletal morphology. It is unclear, however, how much of the variation in WHR is explained by pelvic size (it seems that most of the variance in WHR is due to fat storage on the hip and waist regions).

Effect on the Man's Reproductive Success

The size of the pelvis determines the size of the bony pelvic canal through which the fetus passes during a delivery. As such, a wider pelvis reduces the risk of obstructed labor (Caldwell and Moloy, 1933; Stålberg et al., 2006). In the absence of healthcare, women who are unable to deliver their babies perish, along with their babies. Moreover, obstructed labor can lead to many long-term health issues on both sides, which can influence future survival and fertility. Thus, a woman's small pelvis will decrease the number of descendants a man can sire with her.

However, a large pelvis can be an obstacle to efficient locomotion (Leutenegger, 1974; Lovejoy, 1988; Ruff, 2017 but see Warrener et al., 2015). A woman with a lower ability to walk will have higher difficulties to secure resources for her children, which will decrease their survival or quality. Altogether, stabilizing selection is expected to be operating on female hip size, as well as on men's preferences for this trait.

To conclude, the evolutionary costs and benefits of a wide pelvis seem more appropriate to explain the origin of the sexually dimorphic hip size via natural selection, than to explain men's preferences for a specific WHR. Female pelvic size and shape are the result of two conflicting evolutionary pressures: bipedal locomotion and parturition of a highly encephalized fetus (Leutenegger, 1974; Lovejoy, 1988; Rosenberg and Trevathan, 1995 but see Leong, 2006; Betti and Manica, 2018). It is possible that the link between pelvic size and childbirth and locomotion contributed to the selection of men's preference for an average hip size, but more research is needed to confirm its effect on men's reproductive success.

Perception of the Characteristic Using WHR

To my knowledge, nobody has tested the effect of WHR on perceived difficulties during childbirth, or on perceived locomotion.

Cue of Center of Body Mass

This hypothesis, suggested by Pawłowski and Dunbar (2001) and Pawłowski and Grabarczyk (2003) and found in 6% of the papers

in the literature, stipulates that WHR is linked to the position of the body's center of gravity, which influences bipedal stability.

Correlation With WHR

Everything else being equal, a lower WHR will lower the center of mass of the body. One study uses body measurements of young women to experimentally establish the correlation between WHR and the center of body mass (Pawłowski and Grabarczyk, 2003). However, the correlation is not very strong in their sample of students, and more data is required.

Effect on the Man's Reproductive Success

In advanced pregnancy and during lactation, when the infant is being carried, a bipedal female has to contend with a substantial increase in the anterior load above the center of gravity (Pawłowski, 2001). Fat deposits in the buttocks and thighs may prevent the center of gravity from moving upwards and forwards, and facilitate walking and foraging during pregnancy and lactation. Choosing a mate with a lower center of gravity could increase the survival of the fetus and infants a man would sire with this woman, as she would be less likely to fall and injure the fetus, the infant or herself, and she would be more successful in foraging or escaping predation during these critical periods. A lower center of gravity would also mean a lower energetic cost to maintain balance, and thus an increase in resources available to be directed toward the descendants. Thus, a woman's center of gravity could have an effect on her mate's reproductive success (Pawłowski and Dunbar, 2001; Pawłowski, 2003).

However, as with the pelvic size argument, this hypothesis seems more suitable to explain the origin of dimorphic body shapes in the human species than to explain men's preferences.

Perception of the Characteristic Using WHR

To my knowledge, there has been no research concerning WHR and perceived center of body mass, or perceived walking abilities during pregnancy and lactation.

Cue of Ability to Cope With Stress

The link between stress and women's WHR exists in the literature since 1995 (Singh, 1995, see **Figure 2**), but is included in only 5% of the papers. Depending on the author, a high WHR could be a sign of exogenous stress, a cue of a poor ability to cope with stress, or a cue of an effective response to stress.

Correlation With WHR

Compared to women with low WHRs, women with high WHRs report more chronic stress and have more psychological and psychiatric issues (Björntorp, 1987b, 1993). According to Björntorp, a high WHR might be interpreted as a sign of an inability to cope with environmental stress. One experiment shows that women with high WHRs evaluate laboratory challenges as more threatening, performed more poorly on them, and reported more chronic stress (Epel et al., 2000).

However, Cashdan draws an opposite conclusion from the same observations (Cashdan, 2008). Cortisol (the levels of which are associated with WHR) enables the mind and body to respond effectively to stress, by shifting energy substrates from storage sites to the bloodstream and by increasing blood

pressure and cardiac output. As part of this response, cortisol increases WHR by increasing visceral fat. Conversely, stress-induced cortisol secretion is greater among women with more central fat (Epel et al., 2000).

To conclude, WHR seems to be related to stress responses, but it is not clear if a low WHR is a cue of a good or a poor ability to cope with environmental stress. The stress responses in women with high WHRs may be maladaptive in most WEIRD populations, yet it could be adaptive where conditions are extreme or where stress is episodic rather than constant (Cashdan, 2008).

Effect on the Man's Reproductive Success

If a high WHR is a sign of inadequate coping with stress, women with a high WHR may bear descendants of lower quality because they may be less able to secure resources or provide care for them. However, the opposite is true if a high WHR is a sign of a better ability to respond to stress.

Maternal stress during fetal growth can lead to a lower birthweight. Stress also has epigenetic effects on offspring' life history trajectories and health (Worthman and Kuzara, 2005). However, according to the adaptationist life history perspective, these effects could be associated with a phenotype adapted to the environment (Bateson et al., 2004; Worthman and Kuzara, 2005; Nettle, 2010, see section Cue of Quantity and Availability of "Reproductive Fat").

To conclude, it is unclear if choosing a woman with a lower WHR, as a cue of stress responses, would have a positive, neutral or negative impact on a man's reproductive success. The answer will probably differ according to the environment, and could lead to a preference for a relatively high WHR in some cases (Cashdan, 2008).

Overall, this hypothesis lacks clarity. Nevertheless, the link between stress and WHR is a valuable explanation of the variability of women's WHRs (Cashdan, 2008).

Perception of the Characteristic Using WHR

To my knowledge, the effect of WHR on perceived stress, or ability to cope with stress, has not been investigated.

Cue of Ability to Acquire Resources

It has been suggested that a preference for a relatively high WHR could be adaptive in some environments because the hormonal profile associated with high WHRs (high androgen and cortisol, low estrogen) may favor success in resource competition, particularly under stressful and difficult circumstances (Cashdan, 2008). This hypothesis is mentioned in 10% of the papers.

Correlation With WHR

High androgen levels in women are associated both with higher WHR (Evans et al., 1983; Elbers et al., 1997; Santoro et al., 2005; van Anders and Hampson, 2005) and with greater assertiveness, competitiveness and aggressiveness in women (Purifoy and Koopmans, 1979; Baucom et al., 1985; Dabbs et al., 1988; Cashdan, 1995; Udry et al., 1995; Harris et al., 1996; Dabbs and Hargrove, 1997; Grant and France, 2001; von der Pahlen et al., 2002). Androgens also increase muscle mass and physical

strength (Bhasin et al., 1996). Unfortunately, these studies have been conducted in western countries only, limiting the generalization of the results to other populations.

Androgens also shape features other than WHR (including facial traits, body features and voice; Abitbol et al., 1999; Rickenlund et al., 2003; Lefevre et al., 2013; Whitehouse et al., 2015), and individuals can rely on other cues to assess androgen levels. More importantly, men could use more direct cues of the ability to access resources (e.g., behavior, physical accomplishments or quantity of resources) and may not need indirect cues.

Effect on the Man's Reproductive Success

According to this hypothesis, having a relatively high WHR can increase a woman's survival and reproductive success, because she will be more able to work hard to support herself and her children, compete directly for resources for them, and cope with resource scarcity. Most of these effects will translate into positive effects on her mate's reproductive success.

In this case, the optimal female WHR (for herself and her mate) is likely to vary with the circumstances. In societies where women are expected to provide most of the food, through hard physical work and competition, the balance should be tipped toward a hormonal profile consistent with a higher WHR. In more benign conditions, where women get most of their resources from investing men, a hormonal profile consistent with a low WHR might be more adaptive (Cashdan, 2008).

Overall, as proposed by Cashdan herself, this hypothesis is more likely to explain the variations in women's WHRs (between environments and within lifetime) than to account for men's preferences (Cashdan, 2008). However, we cannot exclude that the link between WHR and women's ability to acquire resources might play a role in the variations observed in the exact value of the preferred WHR between populations.

Perception of the Characteristic Using WHR

One study investigating perceived aggressiveness finds no effect of WHR (Singh, 1994). Another study finds no effect of WHR on factors linked to perceived ambition, independence, self-confidence and success (Henss, 1995). Two studies find that figures with low WHRs are rated as more dominant than figures with high WHRs (Henss, 2000; Buunk and Dijkstra, 2005), which goes in the opposite direction of what is expected according to the present hypothesis. However, these studies are designed to investigate the competition for a mate, not the competition for resources. Studies exploring the effect of WHR on the perceived ability to acquire resources (and not mates) are needed.

Cue of Sex Ratio and Level of Testosterone in Descendants

This hypothesis includes two different sub-hypotheses. The first one, suggested by Manning et al. (1996), stipulates that women with a high WHR have more sons than women with a low WHR, controlling for the total number of children. The second hypothesis states that women with a high WHR have children exhibiting higher levels of testosterone. Pooled together, these two hypotheses are found in 4% of the literature (see **Figure 2**).

Correlation With WHR

A few studies show that a woman's WHR is positively correlated with her number of sons (Manning et al., 1996, 1999; Singh and Zambarano, 1997). However, these studies are measuring women who already have children and correlate WHR with the proportion of existing sons, and it is possible that having sons results in a greater increase in WHR than does having daughters. A more recent study looking at pre-conception WHR and offspring gender finds no significant correlation (Tovée et al., 2001). Thus, there is not enough evidence supporting the fact that a high WHR would be related to more sons in the future.

Manning also found that women with high WHRs tended to have children with low 2D:4D ratios (Manning et al., 1999). A low 2D:4D ratio is supposed to be correlated with high testosterone levels, and the authors conclude that women with high WHRs have more masculine children. However, there is new evidence that the 2D:4D is not a reliable indicator of the levels of testosterone (Hollier et al., 2015; Whitehouse et al., 2015; Apicella et al., 2016).

In conclusion, the idea that a woman with a high WHR will produce more sons or more masculine children is not supported by empirical data.

Effect on the Man's Reproductive Success

Several theories postulate that the sex of the descendants can influence an individual's reproductive success (Hiraiwa-Hasegawa, 1993; Hiraishi et al., 2016). The advantage of sons over daughters depends on various characteristics of both the parents (condition or rank) and the population (including dispersal patterns, inheritance of rank or resources, and degrees of local resource competition). In some cases, one sex has a greater chance of survival and a higher potential reproductive success than the other.

In the hypothetical case where high-WHR women would have children with high testosterone levels, choosing a mate with a relatively high WHR could represent an advantage in some environments. High testosterone is related to various characteristics (from muscular strength to competitive behavior; Bhasin et al., 1996; Apicella et al., 2011, 2015; Schipper, 2014), which could lead to a higher survival and a higher reproductive success.

To conclude, choosing a mate likely to produce more sons or more masculine children could increase the reproductive success of an individual, but it will depend on the environment and on the of the individuals' condition. More importantly, there is no solid evidence that WHR is an indicator of the sex ratio or masculinity of the future descendants. This hypothesis is therefore not supported by empirical evidences.

Perception of the Characteristic Using WHR

The effect of women's WHR on the perception of their children's sex ratio or masculinity has never been investigated.

Cue of Sexual or Maternal Behavior

Interestingly, the idea that a woman's WHR is linked to her behavior and personality, as perceived by others, is found in many of the pioneering papers of this literature (Singh, 1993a,b, 1994; Henss, 1995; Singh and Luis, 1995; Singh and Young, 1995;

Furnham et al., 1998, 2004, 2005; Sugiyama, 2004). However, clear mentions of the hypothesis that WHR could be used as a predictor of past and future behavior by men to choose a mate are rare (2% of the papers) and recent (see Figure 2).

Correlation With WHR

Compared to women with a high WHR, women with a low WHR tend to have a less restricted sociosexuality, sexual intercourse at an earlier age, more sexual partners, and more extrapair copulations (Mikach and Bailey, 1999; Hughes and Gallup, 2003; Fisher et al., 2016). The question remains whether this correlation is due to different preferences and behaviors expressed by women (with hormonal levels as a potential proximal mechanism), or if it only reflects the different opportunities linked to different levels of physical attractiveness. In the latter case, this correlation cannot explain the origin of male preferences for a certain WHR [but it could potentially explain its maintenance, see section Cue of Sexy Daughters (Fisherian Runaway Model)].

Estrogen, testosterone and cortisol levels, all influencing WHR, are linked to maternal investment in many species, including humans (Fleming et al., 1997; Bardi et al., 2001; Dwyer et al., 2004). Thus, WHR could be a cue of women's maternal tendencies. However, there is no direct evidence of a correlation between a woman's WHR before pregnancy and her future maternal investment. Only a few studies provide some indirect evidences for this hypothesis, by showing a correlation between hormonal levels and reported maternal tendencies (Deady and Law Smith, 2006; Deady et al., 2006; Law Smith et al., 2012).

To conclude, more direct evidence is needed to validate the links and mechanisms between women's WHR and their behavior.

Effect on the Man's Reproductive Success

The effect of women's sexual behavior on their mates' reproductive success is double-edged. Women with unrestricted sociosexual orientations, relative to those with more restricted orientations, are more likely to engage in sex at an earlier point in their relationships and have more sexual partners (Simpson and Gangestad, 1991). Thus, being attracted to women with a less restricted sociosexuality might increase the man's chances of mating. On the other hand, women with unrestricted sociosexuality are also more willing to engage in and report higher levels of extradyadic activity (Seal et al., 1994; Barta and Kiene, 2005; Rodrigues et al., 2017; Weiser et al., 2018), therefore increasing the risk of extra-pair copulation costs for their mate (see section Cue of Current Pregnancy). However, these results need to be replicated in non-WEIRD populations before drawing any strong conclusions.

Mating with a woman with a less restricted sociosexuality also increases the risks of being contaminated by sexually transmitted diseases (Hall, 2012). Women with unrestricted sociosexual orientations report more casual sex encounters and multiple and concurrent sexual partners, factors known to increase the risk for exposure to sexually transmitted diseases (Seal and Agostinelli, 1994; Hoyle et al., 2000).

In sum, the effects of a less restricted sociosexuality on the mate's reproductive success are potentially positive for a short-term relationship if the occurrence of sexually transmitted

diseases in the population is low, and probably null or negative for a long-term relationship.

Higher maternal investment can increase the survival and quality of the descendants. However, as stated earlier, to this date there is no direct empirical evidence supporting pre-pregnancy WHR as a cue of future maternal investment.

Overall, this hypothesis has not been explored in many papers, and lacks empirical and theoretical support.

Perception of the Characteristic Using WHR

Several of the early studies investigate the effect of WHR on perceived behavioral and personality traits, but these papers do not include any theoretical background regarding WHR as a potential cue of behavior or personality (Singh, 1993a,b, 1994; Heness, 1995; Singh and Luis, 1995; Singh and Young, 1995; Furnham et al., 1998, 2004, 2005; Sugiyama, 2004). The absence of prediction in these papers is problematic, as the questions asked to the participants are sometimes unclear, and the authors often pooled together items which are linked to different hypotheses, making it impossible to properly test the hypothesis.

Some authors explore the effect of WHR on perceived traits like “*desire for children*,” “*likes children*,” “*good parent*,” or “*nurturing*” (Singh, 1993a,b, 1994; Heness, 1995; Singh and Luis, 1995; Furnham et al., 2005), but the results are inconsistent. Thus, there is no good evidence that WHR is perceived as a cue of maternal behavior, but more appropriate tests with clear predictions are needed.

In a few studies, participants rated figures with high WHRs as more “*faithful*” (Singh, 1994; Singh and Young, 1995). Other studies find that figures with a low WHR are perceived as more “*flirtatious*” (Furnham et al., 2005). These results are in accordance with the hypothesis that WHR serves as a cue of sexual behavior.

Cue of Sexy Daughters (Fisherian Runaway Model)

Fisher famously described a process whereby a small initial preference ultimately leads to extreme traits and preferences through “runaway” selection (Fisher, 1930). If a particular trait in one sex is preferred in mates, then genes disposing stronger preference for the trait could spread as they become linked with genes predisposing the preferred trait.

This hypothesis is not specific to WHR. In fact, the runaway process is almost never applied to men’s preferences for WHR. Yet, in one paper, Singh explains that WHR is heritable and “*offspring of women with lower, more feminine, WHR would have inherited good health and would have been physically attractive to potential mates*” (Singh and Randall, 2007). Tassinary also refers to the runaway model, especially to explain why very small WHRs could theoretically be attractive to men (Tassinary and Hansen, 1998).

Correlation With WHR

For this hypothesis to be valid, WHR needs to be genetically heritable, and there is some evidence that this is the case (Donahue et al., 1992; Bouchard et al., 1996; Schousboe et al., 2004). According to this hypothesis, daughters of women with

a low WHR will have a lower WHR and thus will be more attractive. The hypothesis also requires some heritability of preferences for a low WHR. However, this heritability may cease to be observed once the preference invades the population (since there will not be enough variance in the preferences left). Importantly, this hypothesis does not require any link between WHR and any physiological quality.

Effect on the Man's Reproductive Success

According to this hypothesis, a man mating with a woman with a low WHR will have more attractive daughters than if he mates with a woman with a high WHR. These attractive daughters will have a higher mating and thus reproductive success in the next generation in a population with men attracted by low WHRs, which will have a positive impact on their father’s reproductive success. The size of the effect of women’s WHR on their daughters’ reproductive success remains to be identified. Indirect evidence can be found in studies showing that a low WHR is linked to a higher number of sexual partners, as a proxy for mating success (Mikach and Bailey, 1999; Hughes and Gallup, 2003).

It is important to point out that this hypothesis slightly differs from the other ones in this review because it only involves indirect selection on men’s preferences. A man’s mating preference is favored by direct selection if it increases his own lifetime reproductive output, and by indirect selection if his preference increases the reproductive output of his offspring. Some authors have shown that indirect selection on mate choice via the sexual attractiveness of offspring is a weak evolutionary force relative to direct selection (Kirkpatrick and Barton, 1997). However, such statements of relative strength should not be taken to imply that indirect selection is of little evolutionary importance (Kokko et al., 2003). This would be true only if direct and indirect selections were opposed, which does not seem to be the case for men’s preference for WHR (most of the hypotheses point toward a preference for a low WHR). This hypothesis can then be seen as an additional force reinforcing direct selection on men’s preferences.

Another possible limitation regarding this hypothesis is the indirect cost of sexual antagonism. If WHR is genetically heritable for both sexes, men will have to trade off higher sexiness in daughters with lower-quality sons when choosing a mate, as optimal WHR value differ between men and women (Rice and Chippindale, 2001). Measures of the heritability of WHR for both sexes is necessary to determine the existence of this indirect cost.

Perception of the Characteristic Using WHR

To my knowledge, there is no study investigating the effect of WHR on perceived attractiveness of a woman’s future descendants. The only questions somehow related to this issue are asked by Andrews et al. (2017): “*If this woman were to have a child, it would make friends easily*,” “*If this woman were to have a child, it would be popular*.” They find that the ratings for these items are higher for women with low WHRs. However, these questions were not specifically designed to explore this particular hypothesis.

TABLE 1 | Proposition of classification of the hypotheses found in the literature, according to their theoretical plausibility.

	Unfit hypothesis in light of the current evidence	More evidence is needed to evaluate the plausibility of the hypothesis	Plausible hypothesis for the emergence of men's preferences	Plausible hypothesis for the maintenance of men's preferences	Plausible hypothesis for the selection of women's WHR
1. Cue of biological sex		X	X		
2. Cue of (reproductive) age		X	X		
3. Cue of current pregnancy		X	X		
4. Cue of parity		X	X		
5. Cue of fecundity	X				
6. Cue of quantity and availability of "reproductive fat"		X	X		
7. Cue of health	X				
8. Cue of parasite load	X				
9. Cue of diet	X				
10. Cue of fetal conditions	X				
11. Cue of pelvis size	X			X	
12. Cue of center of body mass				X	
13. Cue of ability to cope with stress	X				
14. Cue of ability to acquire resources			X	X	
15. Cue of sex ratio and level of testosterone in descendants	X				
16. Cue of sexual or maternal behavior	X				
17. Cue of sexy daughters (Fisherian runaway model)				X	

Some hypotheses are unlikely to explain men's preferences toward a certain WHR (column 1). Some hypotheses are likely to explain the emergence of men's preferences during our evolutionary history (column 3), while others better explain the maintenance of these preferences, once they already emerged in the population (column 4). Some hypotheses are good explanations for the selection of a certain WHR in human populations, but do not necessarily lead to a selection of men's preferences (column 5). Finally, some hypotheses need further investigation before one could properly estimate their plausibility (column 2). Both categories and hypotheses are non-exclusive.

Summary of Hypotheses Plausibility

The conclusions of the theoretical analyses of each hypothesis presented in this paper are summarized in **Table 1**. This classification is obviously not definitive and is anticipated to change according to the discovery of new evidence.

DISCUSSION

In this paper, I review the hypotheses explaining why men's preferences for a certain WHR in women may have been selected in the human species. These hypotheses are numerous, and overall, there is some solid theoretical and empirical support in favor of a selection of men's preferences for a mate with a relatively low WHR (with some variations on the exact value according to the population and the environment). However, many of the papers on this topic do not properly develop the theoretical framework, and some interesting hypotheses have been overlooked, while some of the most popular hypotheses require stronger theoretical or empirical support.

To show that men's preference for a certain WHR is an adaptation, it is necessary to demonstrate that a man choosing a mate with a certain WHR will benefit from an increase in reproductive success. Thus, it is crucial to describe the consequences of the preference and show that it can have an impact on the quantity or quality of men's descendants. Importantly, the ultimate focus here is the reproductive success of the individual who is expressing the preference, not of the woman displaying a certain WHR.

WHR as a cue of women's health is one of the most cited hypotheses, appearing in 87% of the papers examined in this review, although health issues linked to WHR have a very limited impact on the women's mates' reproductive success. WHR as a cue of women's fecundity is a notorious hypothesis but is not supported by empirical evidence among populations of young and non-obese women (which is the population of interest for the hypothesis). On the other hand, two hypotheses which are particularly good candidates (WHR as a cue of current pregnancy and parity) are too often forgotten in the literature. Some hypotheses are promising but have been largely overlooked (e.g., WHR as a cue of parasite load, diet or "sexy daughters"). WHR as a cue of quantity and availability of "reproductive fat" hypothesis has received decent empirical and theoretical support and is now generally accepted in the field. WHR as a cue of sex ratio and levels of testosterone in descendants is not supported by empirical evidence, and has therefore never taken hold in the field. Other interesting hypotheses are better suited to explain the presence of a sexually dimorphic WHR in our species through natural selection than men's preferences: WHR as a cue of pelvis size and center of body mass. The preference for slightly higher WHRs in some populations can be explain by WHR as a cue of the ability to acquire resources, although this hypothesis is primarily an excellent account for the variability of women's WHRs. Crucially, the numerous hypotheses reviewed in this paper are not mutually exclusive. The most likely scenario incorporates several of these hypotheses, operating at different periods of our evolutionary history.

To summarize, WHR is a powerful measure (as shown by the numerous physical and physiological characteristics correlated with it), but it may not be as "magical" as often assumed, and not all the features correlated with WHR are linked to mate value. Most of the mate value-related information provided by WHR is relatively basic (sex, age, number of children, current pregnancy). Nevertheless, WHR is a useful and practical visual

trait aggregating the information that a potential mate might not even know is associated with an increase in his own reproductive success.

Non-adaptive explanations for men's preferences toward a certain WHR are not the focus of this paper but they are not necessarily refuted. For example, some authors argue that low WHR preferences may be the result of a generic psychological mechanism of enhanced responding to exaggerated features, or "supernormal" stimuli (Gray et al., 2003). According to this hypothesis, if men view a low WHR as "typical" of female bodies, this could lead men to prefer female WHRs that are even lower than normally attainable (Gray et al., 2003). However, this hypothesis still requires that men use WHR as a cue of biological sex (a hypothesis reviewed in this paper). Men's preferences for a certain WHR can also be explained by sociocultural theories. For example, it is argued that cross-cultural variations in men's preferences for women's WHR could be based on the gender roles occupied by men and women in different cultural settings (Swami et al., 2006a,b). But this hypothesis still requires an explanation regarding the origin of the association between WHR and a certain gender role. Finally, as mentioned earlier, some authors have argued that WHR might not be the best cue of a woman's mate value and that its correlation with attractiveness might be an artifact of men's preferences for another physical characteristic (Tassinari and Hansen, 1998; Tovée et al., 1999; Furnham et al., 2005; Cornelissen et al., 2009a,b; Brooks et al., 2010, 2015; Lassek and Gaulin, 2016). A similar systematic review focusing on a different measure instead of WHR might thus reveal a different picture than the one depicted here (although a few hypotheses concerning men's preferences for features correlated with WHR are incidentally already included in the present review).

The sketch presented by this review calls for more theoretical rigor and precision (and, to be clear, I do include myself in this criticism). Confusion about the theoretical framework can lead to inadequate predictions and suboptimal experimental designs. For example, the stimuli created to test the "WHR as a cue of current pregnancy" hypothesis should be different from the one used to test the "WHR as a cue of (reproductive) age," in terms of WHR range, WHR manipulation (hip or waist changes), and associations with other visual cues (e.g., age of the face). The questions asked to participants to explore the perception of characteristics induced by WHR are often too vague or inadequate, perhaps due to ambiguity in the underlying predictions. The imprecision of the predictions tested previously may have contributed to the increasing number of studies that find null results when testing evolutionary hypotheses for human mating preferences. Null results are not an issue *per se*, but the repeated failure to validate unsound predictions may incorrectly lead to the rejection of an evolutionary explanation to human mate preferences, thus undermining well-founded hypotheses by discrediting the general research paradigm. Finally, the posited theoretical framework will inherently drive the search for the empirical evidence necessary to support a hypothesis. Thus, it is possible that some of the hypotheses presented here would have received more empirical evidences if the theory had been

clearer. For example, most of the evidence used to support the "WHR as a cue of health" hypothesis is not theoretically relevant (health issues at old age or evolutionary recent diseases), maybe in part because of an underspecified theory. With new and more precise predictions, as outlined in this review, additional evidences could be discovered through a deeper exploration of the relevant literatures.

This review has several limitations and should be regarded as a first step to a deeper understanding of this research question, and as a source of ideas to further test the evolutionary origins of mate preferences. I focus only on published research, as the aim is to inventory hypotheses accepted by the academic community, as well as their recognized justifications. However, an examination of unpublished data would be an important next step, in particular to give additional empirical support for the (im)plausibility of the different hypotheses. Moreover, the tentative classification of the different hypotheses presented in **Table 1** is based on their examination through verbal theorizing, but formal models might be helpful to provide a more objective way to define the likelihood of the different scenarios. Quantitative data on the correlations between WHR and the hypothesized traits needs to be gathered, as it would help for the specification of the parameters in such models. Additional layers could be explored to further scrutinize the plausibility of the hypotheses. For example, scenarios where a positive link between women's WHR and her offspring's reproductive success is a necessary condition would require a stronger selection than scenarios based on a higher number and survival of offspring.

In this review, I focus on the literature regarding men's adaptive preferences for women's WHR, but the criticism presented here could be applied to other research topics in evolutionary psychology. It is crucial to establish the evolutionary plausibility of existing hypotheses. Otherwise, we risk hanging on too long to implausible—although often parsimonious—explanations, which can harm the credibility of our field in the long run. Since the replication crisis, much effort has been made to improve our methodological practices, which is extremely encouraging. I hope that this aspiration toward more rigor will also be reflected in how we approach the theoretical foundations of our research.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I thank Vittorio Merola, as well as the three reviewers, for their meticulous and helpful comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01221/full#supplementary-material>

REFERENCES

- Abend, P., Pflüger, L. S., Koppensteiner, M., Coquerelle, M., and Grammer, K. (2015). The sound of female shape: a redundant signal of vocal and facial attractiveness. *Evol. Hum. Behav.* 36, 174–181. doi: 10.1016/j.evolhumbehav.2014.10.004
- Abitbol, J., Abitbol, P., and Abitbol, B. (1999). Sex hormones and the female voice. *J. Voice* 13, 424–446. doi: 10.1016/s0892-1997(99)80048-4
- Amarin, V. N., and Akasheh, H. F. (2001). *Advanced Maternal Age and Pregnancy Outcome*. Available online at: <https://apps.who.int/iris/handle/10665/119069> (accessed March 9, 2019).
- Andrews, T. M., Lukaszewski, A. W., Simmons, Z. L., and Bleske-Rechek, A. (2017). Cue-based estimates of reproductive value explain women's body attractiveness. *Evol. Hum. Behav.* 38, 461–467. doi: 10.1016/j.evolhumbehav.2017.04.002
- Apicella, C. L., Carré, J. M., and Dreber, A. (2015). Testosterone and economic risk taking: a review. *Adapt. Hum. Behav. Physiol.* 1, 358–385. doi: 10.1007/s40750-014-0020-2
- Apicella, C. L., Dreber, A., Gray, P. B., Hoffman, M., Little, A. C., and Campbell, B. C. (2011). Androgens and competitiveness in men. *J. Neurosci. Psychol. Econ.* 4, 54–62. doi: 10.1037/a0021979
- Apicella, C. L., Tobolsky, V. A., Marlowe, F. W., and Miller, K. W. (2016). Hadza hunter-gatherer men do not have more masculine digit ratios (2D:4D). *Am. J. Phys. Anthropol.* 159, 223–232. doi: 10.1002/ajpa.22864
- Bacopoulou, F., Efthymiou, V., Landis, G., Rentoumis, A., and Chrousos, G. P. (2015). Waist circumference, waist-to-hip ratio and waist-to-height ratio reference percentiles for abdominal obesity among Greek adolescents. *BMC Pediatrics* 15:50. doi: 10.1186/s12887-015-0366-z
- Baker, J. L., Olsen, L. W., and Sorensen, T. I. A. (2008). Weight at birth and all-cause mortality in adulthood. *Epidemiology* 19:197. doi: 10.1097/EDE.0b013e31816339c6
- Bardi, M., Shimizu, K., Fujita, S., Borgognini-Tarli, S., and Huffman, M. A. (2001). Hormonal correlates of maternal style in captive macaques (*Macaca fasciata* and *M. mulatta*). *Int. J. Primatol.* 22, 647–662. doi: 10.1023/A:1010793719669
- Barta, W. D., and Kiene, S. M. (2005). Motivations for infidelity in heterosexual dating couples: the roles of gender, personality differences, and sociosexual orientation. *J. Soc. Pers. Relationships* 22, 339–360. doi: 10.1177/0265407505052440
- Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'udine, B., Foley, R. A., et al. (2004). Developmental plasticity and human health. *Nature* 430, 419–421. doi: 10.1038/nature02725
- Baucom, D. H., Besch, P. K., and Callahan, S. (1985). Relation between testosterone concentration, sex role identity, and personality among females. *J. Pers. Soc. Psychol.* 48:1218.
- Beall, C. M., and Goldstein, M. C. (1992). High prevalence of excess fat and central fat patterning among Mongolian pastoral nomads. *Am. J. Hum. Biol.* 4, 747–756. doi: 10.1002/ajhb.1310040606
- Behrman, R. E., Koops, B. L., Morgan, L. J., and Battaglia, F. C. (1982). Neonatal mortality risk in relation to birth weight and gestational age: update. *J. Pediatrics* 101, 969–977. doi: 10.1016/S0022-3476(82)80024-3
- Betti, L., and Manica, A. (2018). Human variation in the shape of the birth canal is significant and geographically structured. *Proc. R. Soc. B* 285:20181807. doi: 10.1098/rspb.2018.1807
- Bhasin, S., Storer, T. W., Berman, N., Callegari, C., Clevenger, B., Phillips, J., et al. (1996). The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N. Engl. J. Med.* 335, 1–7.
- Bjorkelund, C., Lissner, L., Andersson, S., Lapidus, L., and Bengtsson, C. (1996). Reproductive history in relation to relative weight and fat distribution. *Maturitas* 25:78. doi: 10.1016/0378-5122(96)81659-8
- Björntorp, P. (1987a). Fat cell distribution and metabolism. *Ann. N. Y. Acad. Sci.* 499, 66–72. doi: 10.1111/j.1749-6632.1987.tb36198.x
- Björntorp, P. (1987b). The associations between obesity, adipose tissue distribution and disease. *Acta Med. Scand.* 222, 121–134.
- Björntorp, P. (1991). Adipose-tissue distribution and function. *Int. J. Obes.* 15, 67–81.
- Björntorp, P. (1993). Visceral obesity: a “civilization syndrome.” *Obes. Res.* 1, 206–222. doi: 10.1002/j.1550-8528.1993.tb00614.x
- Bleske-Rechek, A., Harris, H. D., Denkinger, K., Webb, R. M., Erickson, L., and Nelson, L. A. (2011). Physical cues of ovulatory status: a failure to replicate enhanced facial attractiveness and reduced waist-to-hip ratio at high fertility. *Evol. Psychol.* 9, 336–353. doi: 10.1177/147470491100900306
- Bohler, H., Mokshagundam, S., and Winters, S. J. (2010). Adipose tissue and reproduction in women. *Fertil. Steril.* 94, 795–825. doi: 10.1016/j.fertnstert.2009.03.079
- Bouchard, C., Rice, T., Lemieux, S., Després, J. P., Périus, L., and Rao, D. C. (1996). Major gene for abdominal visceral fat area in the Québec Family Study. *Int. J. Obes. Relat. Metab. Disord.* 20, 420–427.
- Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: animal communication in a world in flux. *Trends Ecol. Evol.* 25, 292–300. doi: 10.1016/j.tree.2009.11.003
- Brooks, R., Shelly, J. P., Fan, J., Zhai, L., and Chau, D. K. P. (2010). Much more than a ratio: multivariate selection on female bodies. *J. Evol. Biol.* 23, 2238–2248. doi: 10.1111/j.1420-9101.2010.02088.x
- Brooks, R. C., Shelly, J. P., Jordan, L. A., and Dixon, B. J. (2015). The multivariate evolution of female body shape in an artificial digital ecosystem. *Evol. Hum. Behav.* 36, 351–358. doi: 10.1016/j.evolhumbehav.2015.02.001
- Brown, J. E., Potter, J. D., Jacobs, D. R., Kopher, R. A., Rourke, M. J., Barosso, G. M., et al. (1996). Maternal waist-to-hip ratio as a predictor of newborn size: results of the diana project. *Epidemiology* 7, 62–66.
- Butovskaya, M., Sorokowska, A., Karwowski, M., Sabiniewicz, A., Fedenok, J., Dronova, D., et al. (2017). Waist-to-hip ratio, body-mass index, age and number of children in seven traditional societies. *Sci. Rep.* 7:1622. doi: 10.1038/s41598-017-01916-9
- Buunk, B. P., and Dijkstra, P. (2005). A narrow waist versus broad shoulders: sex and age differences in the jealousy-evoking characteristics of a rival's body build. *Pers. Individ. Differ.* 39, 379–389. doi: 10.1016/j.paid.2005.01.020
- Caldwell, W. E., and Moloy, H. C. (1933). Anatomical variations in the female pelvis and their effect in labor with a suggested classification. *Am. J. Obstet. Gynecol.* 26, 479–505. doi: 10.1016/S0002-9378(33)90194-5
- Candolin, U. (2003). The use of multiple cues in mate choice. *Biol. Rev.* 78, 575–595. doi: 10.1017/S1464793103006158
- Casey, V. A., Dwyer, J. T., Berkey, C. S., Bailey, S. M., Coleman, K. A., and Valadian, I. (1994). The distribution of body fat from childhood to adulthood in a longitudinal study population. *Ann. Hum. Biol.* 21, 39–55. doi: 10.1080/03014469400003062
- Cashdan, E. (1995). Hormones, sex, and status in women. *Hormones Behav.* 29, 354–366.
- Cashdan, E. (2008). Waist-to-hip ratio across cultures: trade-offs between androgen- and estrogen-dependent traits. *Curr. Anthropol.* 49, 1099–1107. doi: 10.1086/593036
- Chase, H. C. (1969). Infant mortality and weight at birth: 1960 United States birth cohort. *Am. J. Publ. Health Nations Health* 59, 1618–1628.
- Chen, Z.-P., and Hetzel, B. S. (2010). Cretinism revisited. *Best Pract. Res. Clin. Endocrinol. Metab.* 24, 39–50. doi: 10.1016/j.beem.2009.08.014
- Cornelissen, P. L., Hancock, P. J. B., Kiviniemi, V., George, H. R., and Tovee, M. J. (2009a). Patterns of eye movements when male and female observers judge female attractiveness, body fat and waist-to-hip ratio. *Evol. Hum. Behav.* 30, 417–428. doi: 10.1016/j.evolhumbehav.2009.04.003
- Cornelissen, P. L., Tovee, M. J., and Bateson, M. (2009b). Patterns of subcutaneous fat deposition and the relationship between body mass index and waist-to-hip ratio: implications for models of physical attractiveness. *J. Theoret. Biol.* 256, 343–350. doi: 10.1016/j.jtbi.2008.09.041
- Cross, J. H. (1992). Intestinal capillaritis. *Clin. Microbiol. Rev.* 5, 120–129. doi: 10.1128/CMR.5.2.120
- Dabbs, J., and Hargrove, M. F. (1997). Age, testosterone, and behavior among female prison inmates. *Psychosomat. Med.* 59:477.
- Dabbs, J. M. Jr., Ruback, R. B., Frady, R. L., Hopper, C. H., and Sgoutas, D. S. (1988). Saliva testosterone and criminal violence among women. *Pers. Individ. Differ.* 9, 269–275.
- Deady, D. K., and Law Smith, M. J. (2006). Height in women predicts maternal tendencies and career orientation. *Pers. Individ. Differ.* 40, 17–25. doi: 10.1016/j.paid.2005.06.014
- Deady, D. K., Smith, M. J. L., Sharp, M. A., and Al-Dujaili, E. A. S. (2006). Maternal personality and reproductive ambition in women

- is associated with salivary testosterone levels. *Biol. Psychol.* 71, 29–32. doi: 10.1016/j.biopsycho.2005.01.009
- Donahue, R. P., Prineas, R. J., Gomez, O., and Hong, C. P. (1992). Familial resemblance of body fat distribution: the Minneapolis Children's Blood Pressure Study. *Int. J. Obes. Relat. Metab. Disord.* 16, 161–167.
- Downey, D. B. (2001). Number of siblings and intellectual development: the resource dilution explanation. *Am. Psychol.* 56, 497–504. doi: 10.1037/0003-066X.56.6-7.497
- Dunson, D. B., Colombo, B., and Baird, D. D. (2002). Changes with age in the level and duration of fertility in the menstrual cycle. *Hum. Reprod.* 17, 1399–1403. doi: 10.1093/humrep/17.5.1399
- Dwyer, C. M., Gilbert, C. L., and Lawrence, A. B. (2004). Prepartum plasma estradiol and postpartum cortisol, but not oxytocin, are associated with interindividual and breed differences in the expression of maternal behaviour in sheep. *Hormones Behav.* 46, 529–543. doi: 10.1016/j.yhbeh.2004.05.011
- Eaton, S. B., Eaton, S. B. III, and Konner, M. J. (1997). Review paleolithic nutrition revisited: a twelve-year retrospective on its nature and implications. *Eur. J. Clin. Nutr.* 51:207.
- Eijkemans, M. J. C., Imani, B., Mulders, A. G. M. G. J., Habbema, J. D. F., and Fauser, B. C. J. M. (2003). High singleton live birth rate following classical ovulation induction in normogonadotrophic anovulatory infertility (WHO 2). *Hum. Reprod.* 18, 2357–2362. doi: 10.1093/humrep/deg459
- Elbers, J. M., Asscheman, H., Seidell, J. C., Megens, J. A., and Gooren, L. J. (1997). Long-term testosterone administration increases visceral fat in female to male transsexuals. *J. Clin. Endocrinol. Metab.* 82, 2044–2047.
- Epel, E. S., McEwen, B., Seeman, T., Matthews, K., Castellazzo, G., Brownell, K. D., et al. (2000). Stress and body shape: stress-induced cortisol secretion is consistently greater among women with central fat. *Psychosomat. Med.* 62, 623–632. doi: 10.1097/00006842-200009000-00005
- Evans, D. J., Hoffmann, R. G., Kalkhoff, R. K., and Kissebah, A. H. (1983). Relationship of androgenic activity to body fat topography, fat cell morphology, and metabolic aberrations in premenopausal women. *J. Clin. Endocrinol. Metab.* 57, 304–310. doi: 10.1210/jcem-57-2-304
- Fisher, C. I., Hahn, A. C., DeBruine, L. M., and Jones, B. C. (2016). Is women's sociosexual orientation related to their physical attractiveness? *Pers. Individ. Differ.* 101, 396–399. doi: 10.1016/j.paid.2016.06.022
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection: A Complete Variorum Edition*. Oxford: Oxford University Press.
- Fleming, A. S., Ruble, D., Krieger, H., and Wong, P. Y. (1997). Hormonal and experiential correlates of maternal responsiveness during pregnancy and the puerperium in human mothers. *Hormones Behav.* 31, 145–158. doi: 10.1006/hbeh.1997.1376
- Fredriks, A. M., van Buuren, S., Fekkes, M., Verloove-Vanhorick, S. P., and Wit, J. M. (2005). Are age references for waist circumference, hip circumference and waist-hip ratio in Dutch children useful in clinical practice? *Eur. J. Pediatr.* 164, 216–222. doi: 10.1007/s00431-004-1586-7
- Fretts, R. C., Schmittiel, J., McLean, F. H., Usher, R. H., and Goldman, M. B. (1995). Increased maternal age and the risk of fetal death. *N. Engl. J. Med.* 333, 953–957. doi: 10.1056/NEJM199510123331501
- Furnham, A., Dias, M., and McClelland, A. (1998). The role of body weight, waist-to-hip ratio, and breast size in judgments of female attractiveness. *Sex Roles* 39, 311–326.
- Furnham, A., Lavancy, M., and McClelland, A. (2001). Waist to hip ratio and facial attractiveness: a pilot study. *Pers. Individ. Differ.* 30, 491–502. doi: 10.1016/S0191-8869(00)00040-4
- Furnham, A., McClelland, A., and Omer, L. (2003). A cross-cultural comparison of ratings of perceived fecundity and sexual attractiveness as a function of body weight and waist-to-hip ratio. *Psychol. Health Med.* 8, 219–2304. doi: 10.1080/1354850031000087609
- Furnham, A., Mistry, D., and McClelland, A. (2004). The influence of age of the face and the waist to hip ratio on judgements of female attractiveness and traits. *Pers. Individ. Differ.* 36, 1171–1185. doi: 10.1016/S0191-8869(03)00209-5
- Furnham, A., Moutafi, J., and Baguma, P. (2002). A cross-cultural study on the role of weight and waist-to-hip ratio on female attractiveness. *Pers. Individ. Differ.* 32, 729–745. doi: 10.1016/S0191-8869(01)00073-3
- Furnham, A., Petrides, K. V., and Constantinides, A. (2005). The effects of body mass index and waist-to-hip ratio on ratings of female attractiveness, fecundity, and health. *Pers. Individ. Differ.* 38, 1823–1834. doi: 10.1016/j.paid.2004.11.011
- Furnham, A., Swami, V., and Shah, K. (2006). Body weight, waist-to-hip ratio and breast size correlates of ratings of attractiveness and health. *Pers. Individ. Differ.* 41, 443–454. doi: 10.1016/j.paid.2006.02.007
- Furnham, A., Tan, T., and McManus, C. (1997). Waist-to-hip ratio and preferences for body shape: a replication and extension. *Pers. Individ. Differ.* 22, 539–549.
- Golden, M. H. (1982). Protein deficiency, energy deficiency, and the oedema of malnutrition. *Lancet* 319, 1261–1265. doi: 10.1016/S0140-6736(82)92839-2
- Grant, V. J., and France, J. T. (2001). Dominance and testosterone in women. *Biol. Psychol.* 58, 41–47. doi: 10.1016/S0301-0511(01)00100-4
- Gray, R. D., Heaney, M., and Fairhall, S. (2003). “Evolutionary psychology and the challenge of adaptive explanation,” in *From Mating to Mentality: Evaluating Evolutionary Psychology*, eds K. Sterelny and J. Fitness (New York, NY: Hove: Psychology Press), 247–268.
- Groop, L. (2000). Genetics of the metabolic syndrome. *Br. J. Nutr.* 83, S39–S48. doi: 10.1017/S0007114500000945
- Hackman, E., Emanuel, I., Belle, G., van, and Daling, J. (1983). Maternal birth weight and subsequent pregnancy outcome. *JAMA* 250, 2016–2019. doi: 10.1001/jama.1983.03340150058027
- Hall, N. M. (2012). Sociosexuality, human immunodeficiency virus (HIV) susceptibility, and sexual behavior among African American women. *J. AIDS HIV Res.* 5, 43–51. doi: 10.5897/JAHR12.049
- Harris, J. A., Rushton, J. P., Hampson, E., and Jackson, D. N. (1996). Salivary testosterone and self-report aggressive and pro-social personality characteristics in men and women. *Aggressive Behav.* 22, 321–331.
- Hartz, A. J., Rupley, D. C., and Rimm, A. A. (1984). The association of girth measurements with disease in 32,856 women. *Am. J. Epidemiol.* 119, 71–80.
- Heness, R. (1995). Waist-to-hip ratio and attractiveness. Replication and extension. *Pers. Individ. Differ.* 19, 479–488.
- Heness, R. (2000). Waist-to-hip ratio and female attractiveness. Evidence from photographic stimuli and methodological considerations. *Pers. Individ. Differ.* 28, 501–513. doi: 10.1016/S0191-8869(99)00115-4
- Hernández-Julián, R., Mansour, H., and Peters, C. (2014). The effects of intrauterine malnutrition on birth and fertility outcomes: evidence from the 1974 Bangladesh famine. *Demography* 51, 1775–1796. doi: 10.1007/s13524-014-0326-5
- Hinkle, S. N., Albert, P. S., Mendola, P., Sjaarda, L. A., Yeung, E., Boghossian, N. S., et al. (2014). The association between parity and birthweight in a longitudinal consecutive pregnancy cohort. *Paediatr. Perinatal Epidemiol.* 28, 106–115. doi: 10.1111/ppe.12099
- Hiraishi, K., Murasaki, K., Okuda, H., and Yamate, M. (2016). Sexual and romantic overperception among a Japanese young sample: a replication of Haselton (2003). *Lett. Evol. Behav. Sci.* 7, 29–32. doi: 10.5178/lebs.2016.47
- Hiraiwa-Hasegawa, M. (1993). Skewed birth sex ratios in primates: should high-ranking mothers have daughters or sons? *Trends Ecol. Evol.* 8, 395–400. doi: 10.1016/0169-5347(93)90040-V
- Hollier, L. P., Keelan, J., Jamnadas, E., Maybery, M., Hickey, M., and Whitehouse, A. (2015). Adult digit ratio (2D:4D) is not related to umbilical cord androgen or estrogen concentrations, their ratios or net bioactivity. *Early Hum. Dev.* 91, 111–117. doi: 10.1016/j.earlhumdev.2014.12.011
- Hoyle, R. H., Fejfar, M. C., and Miller, J. D. (2000). Personality and sexual risk taking: a quantitative review. *J. Pers.* 68, 1203–1231. doi: 10.1111/1467-6494.00132
- Hughes, S. M., and Gallup, G. G. (2003). Sex differences in morphological predictors of sexual behavior—shoulder to hip and waist to hip ratios. *Evol. Hum. Behav.* 24, 173–178. doi: 10.1016/S1090-5138(02)00149-6
- Imani, B., Eijkemans, M. J. C., te Velde, E. R., Habbema, J. D. F., and Fauser, B. C. J. M. (2002). A nomogram to predict the probability of live birth after clomiphene citrate induction of ovulation in normogonadotrophic oligoamenorrheic infertility. *Fertil. Steril.* 77, 91–97. doi: 10.1016/S0015-0282(01)02929-6
- Iwasa, Y., and Pomiankowski, A. (1994). The evolution of mate preferences for multiple sexual ornaments. *Evolution* 48, 853–867. doi: 10.1111/j.1558-5646.1994.tb01367.x
- Jasienska, G., Ziolkiewicz, A., Ellison, P. T., Lipson, S. F., and Thune, I. (2004). Large breasts and narrow waists indicate high reproductive potential in women. *Proc. R. Soc. B* 271, 1213–1217. doi: 10.1098/rspb.2004.2712

- Jenkins, J. M., Brook, P. F., Sargeant, S., and Cooke, I. D. (1995). Endocervical mucus pH is inversely related to serum androgen levels and waist to hip ratio. *Fertil. Steril.* 63, 1005–1008. doi: 10.1016/S0015-0282(16)57538-4
- Johnson, K. L., Lurye, L. E., and Tassinari, L. G. (2010). Sex categorization among preschool children: increasing utilization of sexually dimorphic cues. *Child Dev.* 81, 1346–1355. doi: 10.1111/j.1467-8624.2010.01476.x
- Johnson, K. L., and Tassinari, L. G. (2005). Perceiving sex directly and indirectly: meaning in motion and morphology. *Psychol. Sci.* 16, 890–897. doi: 10.1111/j.1467-9280.2005.01633.x
- Johnstone, R. A. (1996). Multiple displays in animal communication: 'backup signals' and 'multiple messages.' *Philos. Transac. R. Soc. Lond. Ser. B Biol. Sci.* 351, 329–338. doi: 10.1098/rstb.1996.0026
- Kaye, S. A., Folsom, A. R., Prineas, R. J., Potter, J. D., and Gapstur, S. M. (1990). The association of body fat distribution with lifestyle and reproductive factors in a population study of postmenopausal women. *Int. J. Obes.* 14, 583–591.
- Kiely, J. L., Paneth, N., and Susser, M. (1986). An assessment of the effects of maternal age and parity in different components of perinatal mortality. *Am. J. Epidemiol.* 123, 444–454. doi: 10.1093/oxfordjournals.aje.a114259
- Kirchengast, S., and Gartner, M. (2002). Changes in fat distribution (WHR) and body weight across the menstrual cycle. *Coll. Antropol.* 26, 47–57.
- Kirkpatrick, M., and Barton, N. H. (1997). The strength of indirect selection on female mating preferences. *PNAS* 94, 1282–1286. doi: 10.1073/pnas.94.4.1282
- Kirschner, M. A., and Samojlik, E. (1991). Sex hormone metabolism in upper and lower body obesity. *Int. J. Obes.* 15 (Suppl. 2), 101–108.
- Kokko, H., Brooks, R., Jennions, M. D., and Morley, J. (2003). The evolution of mate choice and mating biases. *Proc. R. Soc. B* 270, 653–664. doi: 10.1098/rspb.2002.2235
- Kucik, C. J., Martin, G. L., and Sortor, B. V. (2004). Common intestinal parasites. *Am. Family Phys.* 69:2.
- Lanska, D. J., Lanska, M. J., Hartz, A. J., and Rimm, A. A. (1985). Factors influencing anatomic location of fat tissue in 52,953 women. *Int. J. Obes.* 9, 29–38.
- Lassek, W. D., and Gaulin, S. J. (2008). Waist-hip ratio and cognitive ability: is gluteofemoral fat a privileged store of neurodevelopmental resources? *Evol. Hum. Behav.* 29, 26–34. doi: 10.1016/j.evolhumbehav.2007.07.005
- Lassek, W. D., and Gaulin, S. J. C. (2006). Changes in body fat distribution in relation to parity in American women: a covert form of maternal depletion. *Am. J. Phys. Anthropol.* 131, 295–302. doi: 10.1002/ajpa.20394
- Lassek, W. D., and Gaulin, S. J. C. (2016). What makes jessica rabbit sexy? Contrasting roles of waist and hip size. *Evol. Psychol.* 14:59. doi: 10.1177/1474704916643459
- Lassek, W. D., and Gaulin, S. J. C. (2018a). Do the low WHRs and BMIs judged most attractive indicate better health? *Evl. Psychol.* 16:3998. doi: 10.1177/1474704918803998
- Lassek, W. D., and Gaulin, S. J. C. (2018b). Do the Low WHRs and BMIs Judged Most Attractive Indicate Higher Fertility? *Evol. Psychol.* 16:63. doi: 10.1177/1474704918800063
- Law Smith, M. J., Deady, D. K., Moore, F. R., Jones, B. C., Cornwell, R. E., Stirrat, M., et al. (2012). Maternal tendencies in women are associated with estrogen levels and facial femininity. *Horm. Behav.* 61, 12–16. doi: 10.1016/j.yhbeh.2011.09.005
- Law, C. M., Barker, D. J., Osmond, C., Fall, C. H., and Simmonds, S. J. (1992). Early growth and abdominal fatness in adult life. *J. Epidemiol. Community Health* 46, 184–186. doi: 10.1136/jech.46.3.184
- Lefevre, C. E., Lewis, G. J., Perrett, D. I., and Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evol. Hum. Behav.* 34, 273–279. doi: 10.1016/j.evolhumbehav.2013.03.005
- Leibel, R. L., Edens, N. K., and Fried, S. K. (1989). Physiologic basis for the control of body fat distribution in humans. *Ann. Rev. Nutr.* 9, 417–443. doi: 10.1146/annurev.nu.09.070189.002221
- Leong, A. (2006). Sexual dimorphism of the pelvic architecture: a struggling response to destructive and parsimonious forces by natural & mate selection. *Mcgill J. Med.* 9, 61–66.
- Leutenegger, W. (1974). Functional aspects of pelvic morphology in simian Primates. *J. Hum. Evol.* 3, 207–222. doi: 10.1016/0047-2484(74)90179-1
- Ley, C. J., Lees, B., and Stevenson, J. C. (1992). Sex- and menopause-associated changes in body-fat distribution. *Am. J. Clin. Nutr.* 55, 950–954. doi: 10.1093/ajcn/55.5.950
- Lovejoy, C. O. (1988). Evolution of human walking. *Sci. Am.* 259, 118–125.
- Manning, J. T., Anderton, R., and Washington, S. M. (1996). Women's waists and the sex ratio of their progeny: evolutionary aspects of the ideal female body shape. *J. Hum. Evol.* 31, 41–47. doi: 10.1006/jhev.1996.0047
- Manning, J. T., Trivers, R. L., Singh, D., and Thornhill, R. (1999). The mystery of female beauty. *Nature* 399, 214–215. doi: 10.1038/20341
- Manolopoulos, K. N., Karpe, F., and Frayn, K. N. (2010). Gluteofemoral body fat as a determinant of metabolic health. *Int. J. Obes.* 34, 949–959. doi: 10.1038/ijo.2009.286
- Marlowe, F. W., Apicella, C., and Reed, D. (2005). Men's preferences for women's profile waist-to-hip ratio in two societies. *Evol. Hum. Behav.* 26, 458–468. doi: 10.1016/j.evolhumbehav.2005.07.005
- Marlowe, F. W., and Wetsman, A. (2001). Preferred waist-to-hip ratio and ecology. *Pers. Individ. Differ.* 30, 481–489. doi: 10.1016/S0191-8869(00)00039-8
- Marti, B., Tuomilehto, J., Salomaa, V., Kartovaara, L., Korhonen, H. J., and Pietinen, P. (1991). Body fat distribution in the Finnish population: environmental determinants and predictive power for cardiovascular risk factor levels. *J. Epidemiol. Community Health* 45, 131–137. doi: 10.1136/jech.45.2.131
- McCormick, M. C. (1985). The contribution of low birth weight to infant mortality and childhood morbidity. *N. Engl. J. Med.* 312, 82–90.
- McKinnon, C. J., Hatch, E. E., Rothman, K. J., Mikkelsen, E. M., Wesselink, A. K., Hahn, K. A., et al. (2016). Body mass index, physical activity and fecundability in a North American preconception cohort study. *Fertil. Steril.* 106, 451–459. doi: 10.1016/j.fertnstert.2016.04.011
- Menken, J., and Larsen, U. (1986). "Fertility rates and aging," in *Aging, Reproduction, and the Climacteric*, eds L. Mastroianni Jr. and C. A. Paulsen (Oxford: Springer), 147–166. doi: 10.1007/978-1-4684-5047-7_9
- Merklinger-Gruchala, A., Jasienska, G., and Kapiszewska, M. (2015). Short interpregnancy interval and low birth weight: a role of parity. *Am. J. Hum. Biol.* 27, 660–666. doi: 10.1002/ajhb.22708
- Merklinger-Gruchala, A., Jasienska, G., and Kapiszewska, M. (2017). Parity conditions the risk for low birth weight after maternal exposure to air pollution. *Biodemography Soc. Biol.* 63, 71–86. doi: 10.1080/19485565.2016.1264872
- Mikach, S. M., and Bailey, J. M. (1999). What distinguishes women with unusually high numbers of sex partners? *Evol. Hum. Behav.* 20, 141–150. doi: 10.1016/S1090-5138(98)00045-2
- Moller, A. P., and Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behav. Ecol. Sociobiol.* 32, 167–176. doi: 10.1007/BF00173774
- Moran, C., Hernandez, E., Ruiz, J. E., Fonseca, M. E., Bermudez, J. A., and Zarate, A. (1999). Upper body obesity and hyperinsulinemia are associated with anovulation. *Gynecol. Obstet. Invest.* 47, 1–5. doi: 10.1159/000010052
- Mosley, W. H. (1977). "The effects of nutrition on natural fertility," in *Proceedings of a Seminar on Natural Fertility*, eds H. Leridon, J. Menken (Paris: Ordina Editions), 83–105.
- Naeye, R. L. (1983). Maternal age, obstetric complications, and the outcome of pregnancy. *Obstet. Gynecol.* 61, 210–216.
- Nenko, I., and Jasienska, G. (2009). Fertility, body size, and shape: an empirical test of the covert maternal depletion hypothesis. *Am. J. Hum. Biol.* 21, 520–523. doi: 10.1002/ajhb.20938
- Nettle, D. (2010). Dying young and living fast: variation in life history across English neighborhoods. *Behav. Ecol.* 21, 387–395. doi: 10.1093/beheco/arp202
- Osteria, T. S. (1982). Maternal nutrition, infant health, and subsequent fertility. *Philipp J. Nutr.* 35, 106–111.
- Pall, M., Stephens, K., and Azziz, R. (2006). Family size in women with polycystic ovary syndrome. *Fertil. Steril.* 85, 1837–1839. doi: 10.1016/j.fertnstert.2005.11.051
- Pasquali, R., Gambineri, A., and Pagotto, U. (2006). Review article: the impact of obesity on reproduction in women with polycystic ovary syndrome. *Int. J. Obstet. Gynaecol.* 1148–1159. doi: 10.1111/j.1471-0528.2006.00990.x

- Pawlowski, B. (2001). The evolution of gluteal/femoral fat deposits and balance during pregnancy in bipedal homo. *Curr. Anthropol.* 42, 572–574. doi: 10.1086/322548
- Pawlowski, B. (2003). Variable preferences for sexual dimorphism in height as a strategy for increasing the pool of potential partners in humans. *Proc. R. Soc. B* 270, 709–712. doi: 10.1098/rspb.2002.2294
- Pawlowski, B., and Dunbar, R. I. (2001). “Human mate choice strategies,” in *Economics Nature. Social Dilemmas, Mate Choice and Biological Markets*, eds R. Noe, J. van Hoof, and P. Hammerstein (Cambridge: Cambridge University Press), 187–202.
- Pawlowski, B., and Dunbar, R. I. M. (2005). Waist-to-hip ratio versus body mass index as predictors of fitness in women. *Hum. Nat.* 16, 164–177. doi: 10.1007/s12110-005-1002-7
- Pawlowski, B., and Grabarczyk, M. (2003). Center of body mass and the evolution of female body shape. *Am. J. Hum. Biol.* 15, 144–150. doi: 10.1002/ajhb.10136
- Pazhoohi, F., and Liddle, J. R. (2012). Identifying feminine and masculine ranges for waist-to-hip ratio. *J. Soc. Evol. Cult. Psychol.* 6, 227–232. doi: 10.1037/h0099212
- Purifoy, F. E., and Koopmans, L. H. (1979). Androstenedione, testosterone, and free testosterone concentration in women of various occupations. *Soc. Biol.* 26, 179–188.
- Rebuffé-Scrive, M., Enk, L., Crona, N., Lönnroth, P., Abrahamsson, L., Smith, U., et al. (1985). Fat cell metabolism in different regions in women. Effect of menstrual cycle, pregnancy, and lactation. *J. Clin. Invest.* 75, 1973–1976. doi: 10.1172/JCI111914
- Rice, W. R., and Chippindale, A. K. (2001). Intersexual ontogenetic conflict. *J. Evol. Biol.* 14, 685–693. doi: 10.1046/j.1420-9101.2001.00319.x
- Rickenlund, A., Carlström, K., Ekblom, B., Brismar, T. B., von Schoultz, B., and Hirschberg, A. L. (2003). Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance. *Fertil. Steril.* 79, 947–955. doi: 10.1016/s0015-0282(02)04850-1
- Rimm, A. A., Hartz, A. J., and Fischer, M. E. (1988). A weight shape index for assessing risk of disease in 44,820 women. *J. Clin. Epidemiol.* 41, 459–465. doi: 10.1016/0895-4356(88)90047-9
- Rodrigues, D., Lopes, D., and Pereira, M. (2017). Sociosexuality, commitment, sexual infidelity, and perceptions of infidelity: data from the second love web site. *J. Sex Res.* 54, 241–253. doi: 10.1080/00224499.2016.1145182
- Rodrigues, M. L. C. F., and Costa, T. H. M. D. (2001). Association of the maternal experience and changes in adiposity measured by BMI, waist:hip ratio and percentage body fat in urban Brazilian women. *Br. J. Nutr.* 85, 107–114. doi: 10.1079/BJN2000233
- Rosenberg, K., and Trevathan, W. (1995). Bipedalism and human birth: the obstetrical dilemma revisited. *Evol. Anthropol. Issues News Rev.* 4, 161–168. doi: 10.1002/evan.1360040506
- Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Anim. Behav.* 58, 921–931. doi: 10.1006/anbe.1999.1242
- Ruff, C. (2017). Mechanical constraints on the hominin pelvis and the “obstetrical dilemma.” *Anat. Record* 300, 946–955. doi: 10.1002/ar.23539
- Salem, W., Adler, A., Lee, C., and Smith, G. (2012). Maternal waist to hip ratio is a risk factor for macrosomia. *Int. J. Obstetr. Gynaecol.* 119, 291–297. doi: 10.1111/j.1471-0528.2011.03167.x
- Santoro, N., Torrens, J., Crawford, S., Allsworth, J. E., Finkelstein, J. S., Gold, E. B., et al. (2005). Correlates of circulating androgens in mid-life women: the study of women's health across the nation. *J. Clin. Endocrinol. Metab.* 90, 4836–4845. doi: 10.1210/jc.2004-2063
- Saunders, D. R., Williamson, D. K., and Troje, N. F. (2010). Gaze patterns during perception of direction and gender from biological motion. *J. Vis.* 10, 9–9. doi: 10.1167/10.11.9
- Schipper, B. C. (2014). Sex hormones and competitive bidding. *Management Sci.* 61, 249–266. doi: 10.1287/mnsc.2014.1959
- Schousboe, K., Visscher, P., Erbas, B., Kyvik, K. O., Hopper, J., Henriksen, J. E., et al. (2004). Twin study of genetic and environmental influences on adult body size, shape, and composition. *Int. J. Obes.* 28, 39–48. doi: 10.1038/sj.ijo.0802524
- Schützwahl, A. (2006). Judging female figures: a new methodological approach to male attractiveness judgments of female waist-to-hip ratio. *Biol. Psychol.* 71, 223–229. doi: 10.1016/j.biopsycho.2005.04.005
- Seal, D. W., and Agostinelli, G. (1994). Individual differences associated with high-risk sexual behaviour: implications for intervention programmes. *AIDS Care* 6, 393–397. doi: 10.1080/09540129408258653
- Seal, D. W., Agostinelli, G., and Hannett, C. A. (1994). Extradysadic romantic involvement: moderating effects of sociosexuality and gender. *Sex Roles* 31, 1–22. doi: 10.1007/BF01560274
- Sear, R., Mace, R., and McGregor, I. A. (2000). Maternal grandmothers improve nutritional status and survival of children in rural Gambia. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 267, 1641–1647. doi: 10.1098/rspb.2000.1190
- Seidell, J., Baanders-van, E. H., and Ouwehand, I. (1990). Fat distribution in relation to age, degree of obesity, smoking habits, parity and estrogen use: a cross-sectional study in 11,825 Dutch women participating in the DOM-project. *Int. J. Obes.* 14, 753–761.
- Simpson, J. A., and Gangestad, S. W. (1991). Individual-differences in sociosexuality - evidence for convergent and discriminant validity. *J. Pers. Soc. Psychol.* 60, 870–883.
- Singh, D. (1993a). Adaptive significance of female physical attractiveness—role of waist-to-hip ratio. *J. Pers. Soc. Psychol.* 65, 293–307.
- Singh, D. (1993b). Body shape and women's attractiveness - The critical role of waist-to-hip ratio. *Hum. Nat. Interdiscip. Biosoc. Perspect.* 4, 297–321.
- Singh, D. (1994). Is thin really beautiful and good - Relationship between waist-to-hip ratio (WHR) and female attractiveness. *Pers. Individ. Differ.* 16, 123–132.
- Singh, D. (1995). Female health, attractiveness, and desirability for relationships: role of breast asymmetry and waist-to-hip ratio. *Ethol. Sociobiol.* 16, 465–481.
- Singh, D., and Luis, S. (1995). Ethnic and gender consensus for the effect of waist-to-hip ratio on judgment of women's attractiveness. *Hum. Nat. Interdiscip. Biosoc. Perspect.* 6, 51–65.
- Singh, D., and Randall, P. K. (2007). Beauty is in the eye of the plastic surgeon: waist-hip ratio (WHR) and women's attractiveness. *Pers. Individ. Differ.* 43, 329–340. doi: 10.1016/j.paid.2006.12.003
- Singh, D., and Young, R. K. (1995). Body weight, waist-to-hip ratio, breasts, and hips: role in judgments of female attractiveness and desirability for relationships. *Ethol. Sociobiol.* 16, 483–507.
- Singh, D., and Zambarano, R. J. (1997). Offspring sex ratio in women with android body fat distribution. *Hum. Biol.* 69, 545–556.
- Smith, D. E., Lewis, C. E., Caveny, J. L., Perkins, L. L., Burke, G. L., and Bild, D. E. (1994). Longitudinal changes in adiposity associated with pregnancy: the CARDIA study. *JAMA* 271, 1747–1751. doi: 10.1001/jama.1994.03510460039030
- Sorokowski, P., Kościński, K., Sorokowska, A., and Huanca, T. (2014). Preference for women's body mass and waist-to-hip ratio in tsimane' men of the bolivian amazon: biological and cultural determinants. *PLoS ONE* 9:e105468. doi: 10.1371/journal.pone.0105468
- Stålberg, K., Bodestedt, Å., Lyrenäs, S., and Axelsson, O. (2006). A narrow pelvic outlet increases the risk for emergency cesarean section. *Acta Obstetr. Gynecol. Scand.* 85, 821–824. doi: 10.1080/00016340600593521
- Streeter, S. A., and McBurney, D. H. (2003). Waist-hip ratio and attractiveness - New evidence and a critique of “a critical test.” *Evol. Hum. Behav.* 24, 88–98. doi: 10.1016/s1090-5138(02)00121-6
- Sugiyama, L. S. (2004). Is beauty in the context-sensitive adaptations of the beholder? Shiwi use of waist-to-hip ratio in assessments of female mate value. *Evol. Hum. Behav.* 25, 51–62. doi: 10.1016/s1090-5138(03)00083-7
- Sugiyama, L. S. (2005). “Physical attractiveness: an adaptationist perspective,” in *The Handbook of Evolutionary Psychology*, ed D. M. Buss (Hoboken, NJ: John Wiley & Sons, Inc.), 292–343.
- Swami, V., Antonakopoulos, N., Tovée, M. J., and Furnham, A. (2006a). A critical test of the waist-to-hip ratio hypothesis of women's physical attractiveness in Britain and Greece. *Sex Roles* 54, 201–211. doi: 10.1007/s11199-006-9338-3
- Swami, V., Caprario, C., Tovée, M. J., and Furnham, A. (2006b). Female physical attractiveness in Britain and Japan: a cross-cultural study. *Eur. J. Personal.* 20, 69–81. doi: 10.1002/per.568
- Swami, V., Jones, J., Einon, D., and Furnham, A. (2009). Men's preferences for women's profile waist-to-hip ratio, breast size, and ethnic group in Britain and South Africa. *Br. J. Psychol.* 100, 313–325. doi: 10.1348/000712608x329525
- Symons, D. (1981). *The Evolution of Human Sexuality*. Oxford; New York, NY: Oxford University Press.

- Taebi, M., Sadat, Z., Saberi, F., and Kalahroudi, M. A. (2015). Early pregnancy waist-to-hip ratio and risk of preeclampsia: a prospective cohort study. *Hypertens. Res.* 38, 80–83. doi: 10.1038/hr.2014.133
- Tassinari, L. G., and Hansen, K. A. (1998). A critical test of the waist-to-hip-ratio hypothesis of female physical attractiveness. *Psychol. Sci.* 9, 150–155.
- Tchernof, A., and Poehlman, E. T. (1998). Effects of the menopause transition on body fatness and body fat distribution. *Obes. Res.* 6, 246–254. doi: 10.1002/j.1550-8528.1998.tb00344.x
- Tonkelaar, I. D., Seidell, J. C., Van, P. N., Baanders-van, E. H., Jacobus, J. H., and Bruning, P. F. (1989). Factors influencing waist/hip ratio in randomly selected pre- and post-menopausal women in the dom-project (preliminary results). *Int. J. Obes.* 13, 817–824.
- Tovée, M. J., Brown, J. E., and Jacobs, D. (2001). Maternal waist-to-hip ratio does not predict child gender. *Proc. R. Soc. B* 268, 1007–1010. doi: 10.1098/rspb.2001.1618
- Tovée, M. J., Furnham, A., and Swami, V. (2007). “Healthy body equals beautiful body? Changing perceptions of health and attractiveness with shifting socioeconomic status,” in *The Body Beautiful*, eds V. Swami and A. Furnham (Oxford: Palgrave Macmillan), 108–128. doi: 10.1057/9780230596887_7
- Tovée, M. J., Hancock, P. J. B., Mahmoodi, S., Singleton, B. R. R., and Cornelissen, P. L. (2002). Human female attractiveness: waveform analysis of body shape. *Proc. R. Soc. B* 269, 2205–2213. doi: 10.1098/rspb.2002.2133
- Tovée, M. J., Maisey, D. S., Emery, J. L., and Cornelissen, P. L. (1999). Visual cues to female physical attractiveness. *Proc. R. Soc. B* 266, 211–218.
- Troisi, R. J., Wolf, A. M., Manson, J. A. E., Klingler, K. M., and Colditz, G. A. (1995). Relation of body fat distribution to reproductive factors in pre- and postmenopausal women. *Obes. Res.* 3, 443–451. doi: 10.1002/j.1550-8528.1995.tb00139.x
- Udry, J. R., Morris, N. M., and Kovenock, J. (1995). Androgen effects on women's gendered behaviour. *J. Biosoc. Sci.* 27, 359–368. doi: 10.1017/S0021932000022884
- van Anders, S. M., and Hampson, E. (2005). Waist-to-hip ratio is positively associated with bioavailable testosterone but negatively associated with sexual desire in healthy premenopausal women. *Psychosomat. Med.* 67:246. doi: 10.1097/01.psy.0000151747.22904.d7
- von der Pahlen, B., Lindman, R., Sarkola, T., Mäkilä, H., and Eriksson, C. J. P. (2002). An exploratory study on self-evaluated aggression and androgens in women. *Aggressive Behav.* 28, 273–280. doi: 10.1002/ab.80005
- Wallace, W. H. B., and Kelsey, T. W. (2010). Human ovarian reserve from conception to the menopause. *PLoS ONE* 5:e8772. doi: 10.1371/journal.pone.0008772
- Wang, G., Djafarian, K., Egedigwe, C. A., El Hamdouchi, A., Ojiambo, R., Ramuth, H., et al. (2015). The relationship of female physical attractiveness to body fatness. *PeerJ* 3:e1155. doi: 10.7717/peerj.1155
- Warren, A. G., Lewton, K. L., Pontzer, H., and Lieberman, D. E. (2015). A wider pelvis does not increase locomotor cost in humans, with implications for the evolution of childbirth. *PLoS ONE* 10:e0118903. doi: 10.1371/journal.pone.0118903
- Wass, P., Waldenström, U., Rossner, S., and Hellberg, D. (1997). An android body fat distribution in females impairs the pregnancy rate of *in-vitro* fertilization-embryo transfer. *Hum. Reprod.* 12, 2057–2060.
- Waterlow, J. C. (1984). Kwashiorkor revisited: the pathogenesis of oedema in kwashiorkor and its significance. *Trans. R. Soc. Trop. Med. Hyg.* 78, 436–441. doi: 10.1016/0035-9203(84)90053-1
- Weinstein, M., Wood, J. W., Stoto, M. A., and Greenfield, D. D. (1990). Components of age-specific fecundability. *Populat. Stud.* 44, 447–467. doi: 10.1080/0032472031000144846
- Weiser, D. A., Niehuis, S., Flora, J., Punyanunt-Carter, N. M., Arias, V. S., and Hannah Baird, R. (2018). Swiping right: sociosexuality, intentions to engage in infidelity, and infidelity experiences on Tinder. *Pers. Individ. Differ.* 133, 29–33. doi: 10.1016/j.paid.2017.10.025
- Wells, J. C. K., Charoensiriwath, S., and Treleaven, P. (2011). Reproduction, aging, and body shape by three-dimensional photonic scanning in Thai men and women. *Am. J. Hum. Biol.* 23, 291–298. doi: 10.1002/ajhb.21151
- Wells, J. C. K., Griffin, L., and Treleaven, P. (2010). Independent changes in female body shape with parity and age: a life-history approach to female adiposity. *Am. J. Hum. Biol.* 22, 456–462. doi: 10.1002/ajhb.21017
- Wenzlaff, F., Briken, P., and Dekker, A. (2018). If there's a penis, it's most likely a man: investigating the social construction of gender using eye tracking. *PLoS ONE* 13:e0193616. doi: 10.1371/journal.pone.0193616
- Wetsman, A., and Marlowe, F. W. (1999). How universal are preferences for female waist-to-hip ratios? Evidence from the Hadza of Tanzania. *Evol. Hum. Behav.* 20, 219–228.
- Whitehouse, A. J. O., Gilani, S. Z., Shafait, F., Mian, A., Tan, D. W., Maybery, M. T., et al. (2015). Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proc. R. Soc. B* 282:20151351. doi: 10.1098/rspb.2015.1351
- Wise, L. A., Palmer, J. R., and Rosenberg, L. (2013). Body size and time-to-pregnancy in black women. *Hum. Reprod.* 28, 2856–2864. doi: 10.1093/humrep/det333
- Wise, L. A., Rothman, K. J., Mikkelsen, E. M., Sørensen, H. T., Riis, A., and Hatch, E. E. (2010). An internet-based prospective study of body size and time-to-pregnancy. *Hum. Reprod.* 25, 253–264. doi: 10.1093/humrep/dep360
- Worthman, C. M., and Kuzara, J. (2005). Life history and the early origins of health differentials. *Am. J. Hum. Biol.* 17, 95–112. doi: 10.1002/ajhb.20096
- Yamamoto, S., Douchi, T., Yoshimitsu, N., Nakae, M., and Nagata, Y. (2001). Waist to hip circumference ratio as a significant predictor of preeclampsia, irrespective of overall adiposity. *J. Obstet. Gynaecol. Res.* 27, 27–31. doi: 10.1111/j.1447-0756.2001.tb01211.x
- Yu, D. W., and Shepard, G. H. Jr. (1998). Is beauty in the eye of the beholder? *Nature* 396, 321–322.
- Zaadstra, B. M., Seidell, J. C., Van Noord, P. A., te Velde, E. R., Habbema, J. D., Vrieswijk, B., et al. (1993). Fat and female fecundity: prospective study of effect of body fat distribution on conception rates. *BMJ* 306, 484–487.
- Zavos, P. M., and Cohen, M. R. (1980). The pH of cervical mucus and the postcoital test. *Fertil. Steril.* 34, 234–238. doi: 10.1016/S0015-0282(16)44953-8

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bovet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Skin Color Preferences in a Malaysian Chinese Population

Kok Wei Tan^{1,2*} and Ian D. Stephen^{3,4,5}

¹School of Psychology and Clinical Language Sciences, University of Reading Malaysia, Iskandar Puteri, Malaysia, ²School of Psychology, University of Nottingham Malaysia, Semenyih, Malaysia, ³Department of Psychology, Macquarie University, North Ryde, NSW, Australia, ⁴ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, North Ryde, NSW, Australia, ⁵Perception in Action Research Centre, Macquarie University, North Ryde, NSW, Australia

OPEN ACCESS

Edited by:

T. Joel Wade,
Bucknell University, United States

Reviewed by:

David Perrett,
University of St Andrews,
United Kingdom
Xue Lei,
University of St Andrews,
United Kingdom
in collaboration with reviewer DP
Justin Kyle Mogilski,
University of South Carolina
Salkehatchie, United States

*Correspondence:

Kok Wei Tan
t.kokwei@reading.edu.my

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 30 November 2018

Accepted: 24 May 2019

Published: 19 June 2019

Citation:

Tan KW and Stephen ID (2019) Skin
Color Preferences in a Malaysian
Chinese Population.
Front. Psychol. 10:1352.
doi: 10.3389/fpsyg.2019.01352

Facial skin color influences the perceived health and attractiveness of Caucasian faces, and has been proposed as a valid cue to aspects of physiological health. Similar preferences for skin color have previously been found in African participants, while different preferences have been found among mainland Chinese participants. Here, we asked Malaysian Chinese participants (ethnic Chinese living in an Asian country with high levels of exposure to Western culture) to manipulate the skin color of Malaysian Chinese, Caucasian, and African faces to make them “look as healthy as possible.” Participants chose to increase skin yellowness to a greater extent than to increase skin redness to optimize healthy appearance. The slight reduction in skin lightness chosen was not statistically significant after correction for multiple comparisons. While broadly in line with the preferences of Caucasian and African participants from previous studies, this differs from mainland Chinese participants. There may be a role for culture in skin color preferences, though methodological differences mean that further research is necessary to identify the cause of these differences in preferences.

Keywords: face perception, skin color, perceived health, Asian, culture difference

INTRODUCTION

Since the 1990s, evolutionary psychologists have theorized that attractiveness and health judgments serve as a mechanism for identifying a healthy, fertile mate (for a review, see Stephen and Tan, 2015). While some facial cues, such as symmetry and averageness, may be perceived as attractive universally across populations (for reviews, see Rhodes, 2006; Little et al., 2011), other cues, such as body size, appear to vary cross-culturally, either due to cultural or ecological differences between populations (Tovée et al., 2006, 2007).

Facial skin color has been shown to influence perceived attractiveness and health (Stephen et al., 2009b, 2012; Whitehead et al., 2012b; Pezdirc et al., 2017), with increased lightness (represented by the L* dimension in CIELab color space), redness (a*), and yellowness (b*) perceived as healthier. Preference for skin redness may be related to increased blood oxygenation (Stephen et al., 2009a), which serves as an indicator for aerobic fitness and fertility (Armstrong and Welsman, 2001; Charkoudian, 2003; Barelli et al., 2007). Skin luminance is determined by concentration of melanin (Edwards and Duntley, 1939) which is associated with health benefits such as photoprotection and synthesis of Vitamin D (Jablonski and Chaplin, 2000; Kourosh et al., 2010). Skin yellowness is influenced by levels of yellow-red carotenoid pigments

in the skin. These carotenoid pigments are obtained from fruit and vegetables in the diet, and then deposited in the top layer of the skin, the *stratum corneum* (Alaluf et al., 2002), and research in Caucasian, Asian, and African populations have linked fruit and vegetable consumption with increased skin yellowness (Stephen et al., 2011; Whitehead et al., 2012a; Tan et al., 2015). Carotenoids are thought to be beneficial for human immunity, visual acuity, and photoprotection of skin (Hughes, 1999; Samimi, 2005; Rao and Rao, 2007; Darvin et al., 2011).

Studies in Caucasian and Asian samples have shown that skin coloration associated with increased intake of fruit and vegetables is perceived as healthy and attractive (Stephen et al., 2011; Whitehead et al., 2012b,c; Lefevre et al., 2013; Lefevre and Perrett, 2015; Pezdirc et al., 2017; Tan et al., 2017), though some studies that did not control for facial expression (Appleton et al., 2018) or color calibration of images and monitors (Jones, 2018) have failed to replicate these results. While similar patterns of preferences for skin lightness, redness, and yellowness coloration have been reported for Caucasian and African samples (Stephen et al., 2009a, 2011; Coetzee et al., 2012), a recent study suggested that mainland Chinese participants show a weaker preference for increased redness and a stronger preference for increased lightness than Caucasian participants and, in contrast to Caucasian and African samples, prefer decreased yellowness (Han et al., 2018).

While the discrepancy between mainland Chinese and other samples may be explained by differences in methodology – Han et al. (2018) used a forced-choice paradigm that did not allow for the elucidation of the amount of color change that was perceived as healthiest – these differences may be attributable to cultural difference in skin color preference. Such cross-cultural differences have been found for other aspects of attractiveness preferences, including preferences for female body size (Swami and Tovée, 2005; Tovée et al., 2006) and male facial masculinity (DeBruine et al., 2010; Brooks et al., 2011), and may relate to cultural differences in cognitive process (Blais et al., 2008; Tan et al., 2012).

Malaysian Chinese, while ethnically Han Chinese, live in a Southeast Asian country that is strongly multicultural (61.7% of the population are ethnic Malay or indigenous, 20.8% Chinese, 6.2% Indian, 0.9% other, and 10.4% noncitizens; CIA, 2018) and influenced by Western culture (86% of movies shown in Malaysian cinemas are Western, 14% local, compared with 56% local movies in China; Epstein, 2011). Previous studies of face perception have found Malaysian Chinese participants to show patterns intermediate between Western and mainland Chinese samples (Tan et al., 2012), and exposure to Western culture has been shown to change individuals' face recognition strategies (Sangrigoli et al., 2005; Hancock and Rhodes, 2008) and attractiveness preferences (Tovée et al., 2006; Boothroyd et al., 2016). Malaysian Chinese participants have been found to show reduced (though still positive) preference for carotenoid coloration, which contains a large b^* component, compared to Western participants in an experimental study (Tan et al., 2017), and to show preferences for lighter and yellower, but not redder, skin in a correlational design (Tan et al., 2018). However, it is not yet known whether Malaysian Chinese show preferences for redness (a^*), yellowness (b^*), and lightness (L^*) in line with Western participants.

Here, we examine Malaysian Chinese participants' preferences for facial skin color by allowing them to manipulate – separately – facial skin lightness (L^*), redness (a^*), and yellowness (b^*) to optimize the healthy appearance of Asian, Caucasian, and African faces. In line with previous studies (Stephen et al., 2009a,b), we predicted that participants will increase skin redness, yellowness, and luminance to enhance the healthy appearance of faces (Lefevre and Perrett, 2015; Pezdirc et al., 2017; Tan et al., 2018). Previous studies have shown reduced preferences for yellowness and redness, and increased preference for lightness in Asian faces (Tan et al., 2017; Han et al., 2018) but not African faces (Stephen et al., 2011), compared to Caucasian faces, as perceived by own-race observers. However, the influence of color on perceptions of attractiveness has also been shown to be reduced in other-race faces, possibly due to unfamiliarity effects (Stephen et al., 2012). We predict a similar pattern of preferences for Malaysian Chinese participants observing Asian and Caucasian faces, and reduced effects in less familiar African faces.

MATERIALS AND METHODS

Stimuli

Twelve facial photographs of three different ethnicities (four Caucasian, four African, and four East Asian) were obtained from Stephen et al. (2017). These photographs were taken under controlled conditions and color calibrated using Psychomorph (Tiddeman et al., 2001). Hair was held back from the face with a black head band, and participants were asked to pose with a neutral expression while holding a Munsell N5 painted board over their shoulders to obscure clothing.

Matlab was used to produce masks with even coloration representing the skin areas of faces, with a Gaussian blur at the edges. One mask was created to represent average face color +8 units of a^* (increased redness) and another one with average face color –8 units of a^* (decreased redness). Color changes are described using CIE $L^*a^*b^*$ color space, in which colors are described along L^* [which takes values between 0 (darkest) and 100 (lightest)], a^* [which takes values between –110 (greenest) and 110 (reddest)], and b^* [which takes values between –110 (bluest) and 110 (yellowest)], which is designed to reflect the way in which the human visual system processes color information, and is perceptually uniform so that a change of 1 unit in one dimension is perceptually equivalent in magnitude to a change of 1 unit in another dimension (Martinkauppi, 2002). The Euclidean distance (ΔE) between two points in CIE $L^*a^*b^*$ space mirrors the color differences as perceived by human vision (Wyszecki and Stiles, 1982). The facial redness of all the 12 faces used was transformed by the difference in color between each of the pairs of masks, in a series of 13 steps. This produced a series of 13 frames, numbered from 0 to 12, whereby frame 0 had skin redness reduced by 8 units of a^* , increasing incrementally so that frame 6 was the original image and frame 12 had skin redness increased by 8 units of a^* . Hair, eyes, clothing, and the background were not manipulated. This procedure was repeated for L^* (lightness) and b^* (yellowness) color axes (Figure 1).



FIGURE 1 | CIELab color transformation of a face with decreased (top) and increased (bottom) lightness (L^* , left); redness (a^* , middle); and yellowness (b^* , right). Presented face is a composite for illustration purposes, but photographs of real individuals were used as the stimuli.

Participants and Procedure

Forty-four Malaysian Chinese participants (18 males, 26 females; mean age = 22.05, $SD = 1.23$) were recruited for this study, giving 95% power to detect small to medium effect sizes in the hypothesized main effects and interactions. All participants were students at Universiti Tunku Abdul Rahman.

Stimuli were presented using computers attached to 15" TFT monitors that were color calibrated with a DataColor Spyder3 Pro. Participants were presented with facial images, one image at a time, and were asked to adjust the color of skin portions of the facial images presented to "make the face look as healthy as possible." By moving the mouse horizontally, the participants cycled through the 13 frames of the transform (same face, different level of color intensity). The participants clicked on the mouse when they felt that the face looked the healthiest.

Each facial image was presented once in each of the three different color dimensions (lightness, redness, yellowness), making a total of 36 trials (12 faces \times 3 color dimensions). The location of the midpoint was randomized and the transform looped to obscure the location of the original facial color, and the order of the trials was also randomized in a single block.

RESULTS

Mean color changes that were applied to the 12 faces along each color axis were calculated. One-sample t -tests showed that participants increased facial yellowness by 1.32 units ($SD = 1.28$), $t(43) = 6.85$, $p < 0.001$, and facial redness by 0.78 units ($SD = 1.09$), $t(43) = 4.72$, $p < 0.001$, and decreased facial lightness by 0.37 units ($SD = 1.06$), $t(43) = 2.29$, $p = 0.027$ (**Figure 2**), to optimize healthy appearance. Comparing these values to typical values of facial lightness (L^*), redness (a^*), and yellowness (b^*) of the studied populations obtained from previous datasets (Stephen et al., 2012; Tan et al., 2018), they correspond to an increase of 0.62 SD for b^* , 0.44 SD for a^* and -0.07 SD for luminance. The result for skin lightness is no longer significant after Bonferroni correction for multiple comparisons.

A 4-way mixed ANOVA was run to examine the differences in the amount of color change applied to the faces of different sexes and ethnicities for all the three color axes, and participants of both genders.

There was a significant main effect for color axes, $F(2, 84) = 25.91$, $p < 0.001$, $\eta_p^2 = 0.38$. Bonferroni-corrected pairwise comparisons showed a significantly greater increment in facial yellowness than redness (mean difference = 0.55, $p = 0.037$) or luminance (mean difference = 1.81, $p < 0.001$),

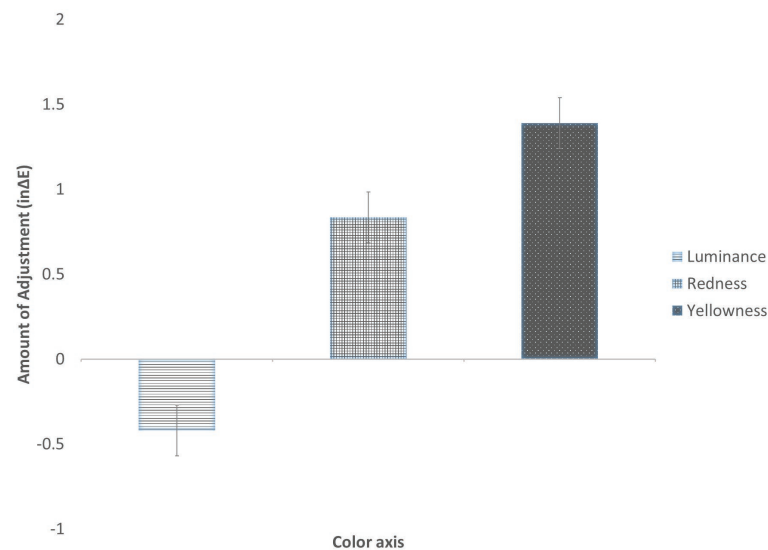


FIGURE 2 | Amount of facial skin color change selected to optimize apparent health of faces.

and greater increment in redness than luminance (mean difference = 1.25, $p < 0.001$).

A significant main effect of ethnicity was found $F(2, 84) = 41.5977$, $p < 0.001$, $\eta_p^2 = 0.50$. Caucasian faces received significantly more positive color adjustment as compared to African faces (mean difference = 1.05, $p < 0.001$) and Asian faces (mean difference = 0.36, $p = 0.004$). Asian faces also received more positive adjustment in facial coloration than African faces (mean difference = 0.69, $p < 0.001$).

No significant main effect was found for sex of face, $F(1, 42) = 0.015$, $p = 0.902$, $\eta_p^2 = 0.000$, nor sex of participants, $F(1, 42) = 2.97$, $p = 0.092$, $\eta_p^2 = 0.066$.

There was a significant interaction for color \times gender $F(2, 84) = 4.357$, $p = 0.016$, $\eta_p^2 = 0.094$. We ran two one-way ANOVAs with repeated measures, and the pairwise comparison showed that, for male participants, there was a significant difference in their adjustment on skin luminance and skin redness (mean difference = -1.88, SE = 0.51, $p = 0.006$) and on skin luminance and skin yellowness (mean difference = -2.50, SE = 0.59, $p = 0.002$). There was no significant difference in adjustment on skin redness and skin yellowness ($p = 0.35$). For female participants, only the adjustment of skin luminance and skin yellowness was significantly different (mean difference = -1.12, SE = 0.26, $p = 0.001$).

There is a significant interaction for ethnic \times gender of face \times gender, $F(2, 84) = 4.106$, $p = 0.02$, $\eta_p^2 = 0.089$, which is out of our main research focus. All the other interactions were not significant ($p > 0.05$).

DISCUSSION

The current study examined Malaysian Chinese participants' perception of healthy facial skin color. Participants increased the skin yellowness to a greater extent, and increased skin redness to a lesser extent to make faces of three ethnicities look as

healthy as possible. The change in skin lightness was not significant after Bonferroni correction. A similar pattern of preferences for skin redness and skin yellowness was observed in previous studies, whereby Caucasian participants significantly increased facial skin yellowness and redness to optimize perceived facial health (Stephen et al., 2009b, 2011; Coetzee et al., 2012; Han et al., 2018). It may be that the observed preference for redder facial skin may be attributable to the appearance of the perfusion of the skin with oxygenated blood, which is associated with physical fitness and increased levels of sex hormones (Armstrong and Welsman, 2001; Charkoudian, 2001; Stephen et al., 2009a). Similarly, increased skin yellowness has been associated with higher levels of deposition of antioxidant carotenoids in the skin, associated with a diet rich in fruits and vegetables (Stephen et al., 2011; Lefevre et al., 2013; Pezdirc et al., 2014, 2017; Tan et al., 2015, 2017). Previous studies also found that preference for skin yellowness was stronger than that of skin redness and skin luminance (Lefevre and Perrett, 2015; Tan et al., 2018), which has been suggested to be related to the antioxidant properties of carotenoids (Paiva and Russell, 1999; Stahl and Sies, 2003), and its protective values to humans' physical health (Hughes, 2001; Tapiero et al., 2004; Krinsky and Johnson, 2005; Samimi, 2005; Rao and Rao, 2007).

However, it should be noted that Han et al. (2018) failed to find preferences for yellowness in mainland Chinese faces using a two-alternative forced-choice (2AFC) paradigm. Preferences of skin redness for mainland Chinese faces were also not as strong as those observed in the Caucasian sample. The manipulations used by Han et al. (2018), however, were more than double the amount of yellowness increment chosen by participants in the current study, and more than double the amount of carotenoid-induced color change preferred by Malaysian Chinese participants in a previous study (Tan et al., 2017), more than 1.5 SD of the yellowness in an Asian population (Tan et al., 2018), and more than triple the amount of color change preferred by participants

in the current study. It may be, therefore, that the high redness and high yellowness images used by Han et al. (2018) were more extreme than looks healthy, and therefore real color preferences may have been obscured.

However, while previous studies have found that Caucasian, African, and mainland Chinese participants choose to increase the lightness of facial skin to optimize healthy appearance (Stephen et al., 2009b; Coetzee et al., 2012; Han et al., 2018), in this study, Malaysian Chinese participants decreased skin lightness, though this preference was no longer significant after Bonferroni correction. While Chinese diaspora culture typically values lighter skin, particularly in women (a common Chinese saying is “a fair skin can hide three facial flaws”; Mak, 2007), this may be offset in the Malaysian context, where ultraviolet radiation from the sun is frequently intense (Kuala Lumpur is less than 400 km from the equator), and increased levels of melanin provides increased protection from sunburn and skin cancer (Jablonski and Chaplin, 2000; Jablonski, 2004).

Previous studies have suggested that skin color changes are more easily detectable in lighter than in darker skinned populations (Coetzee and Perrett, 2014). In the current study, the amount of color adjustment made to optimize the apparent health of faces was greatest for Caucasian faces, followed by Asian, and then African faces, suggesting that skin color may play a greater role in the perception of health in faces from lighter skinned populations.

Limitations

It should be noted that the current paper allowed participants to manipulate the faces along each color axis separately. However, it may be that the color axes interact such that changes in one color axis affect preferences for color on a different axis. Studies in which all color axes are manipulated simultaneously are required to address this question.

REFERENCES

- Alaluf, S., Heinrich, U., Stahl, W., Tronnier, H., and Wiseman, S. (2002). Human nutrition and metabolism: dietary carotenoids contribute to normal human skin color and UV photosensitivity. *J. Nutr.* 132, 399–403. doi: 10.1093/jn/132.3.399
- Appleton, K. M., Mcgrath, A. J., Mckinley, M. C., Draffin, C. R., Hamill, L. L., Young, I. S., et al. (2018). The value of facial attractiveness for encouraging fruit and vegetable consumption: analyses from a randomized controlled trial. *BMC Public Health* 18:298. doi: 10.1186/s12889-018-5202-6
- Armstrong, N., and Welsman, J. (2001). Peak oxygen uptake in relation to growth and maturation in 11- to 17-year-old humans. *Eur. J. Appl. Physiol.* 85, 546–551. doi: 10.1007/s004210100485
- Barelli, C., Heistermann, M., Boesch, C., and Reichard, U. H. (2007). Sexual swellings in wild white-handed gibbon females (*Hylobates lar*) indicate the probability of ovulation. *Horm. Behav.* 51, 221–230. doi: 10.1016/j.yhbeh.2006.10.008
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., and Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One* 3:e3022. doi: 10.1371/journal.pone.0003022
- Boothroyd, L. G., Jucker, J. L., Thornborrow, T., Jamieson, M. A., Burt, D. M., Barton, R. A., et al. (2016). Television exposure predicts body size ideals in rural Nicaragua. *Br. J. Psychol.* 107, 752–767. doi: 10.1111/bjop.12184
- Brooks, R., Scott, I. M., Maklakov, A. A., Kasumovic, M. M., Clark, A. P., and Penton-Voak, I. S. (2011). National income inequality predicts women's

While some discrepancies in skin color preference have been observed across studies conducted at different geographical locations, it cannot be confidently concluded from these data that cultural differences account for the differences between the preferences shown here by Malaysian Chinese participants and participants from Western, African, and mainland Chinese populations. Studies in which methodology is standardized across multiple locations and in which measures of culture are deployed should be conducted to confirm the role of culture, as opposed to methodological differences or ecological differences, in driving the different preferences across populations.

In conclusion, Malaysian Chinese participants show a pattern of facial skin color preference intermediate between that reported in mainland Chinese (Han et al., 2018) and Western (Stephen et al., 2009b) populations, though more similar to the Westerners. While it may be speculated that exposure to Western culture may explain this pattern of results, future studies should standardize methodology across multiple geographical locations, and include measures of culture to confirm this hypothesis.

ETHICS STATEMENT

All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee at the University of Nottingham Malaysia Campus.

AUTHOR CONTRIBUTIONS

KWT and IS contributed to the conception and design of the study. KWT collected the research data. Both authors performed the statistical analysis, wrote the manuscript, and read and approved the submitted version.

- preferences for masculinized faces better than health does. *Proc. Royal Soc. Lond.* 278, 810–812. discussion 813–4. doi: 10.1098/rspb.2010.0964
- Charkoudian, N. (2001). Influences of female reproductive hormones on sympathetic control of the circulation in humans. *Clin. Auton. Res.* 11, 295–301. doi: 10.1007/BF02332974
- Charkoudian, N. (2003). Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. *Mayo Clin. Proc.* 78, 603–612. doi: 10.4065/78.5.603
- CIA. (2018). Available at: <https://www.cia.gov/library/publications/the-world-factbook/geos/my.html> (Accessed January 24, 2018).
- Coetzee, V., Faerber, S. J., Greeff, J. M., Lefevre, C. E., Re, D. E., and Perrett, D. I. (2012). African perceptions of female attractiveness. *PLoS One* 7:e48116. doi: 10.1371/journal.pone.0048116
- Coetzee, V., and Perrett, D. I. (2014). Effect of beta-carotene supplementation on African skin. *J. Biomed. Opt.* 19:25004. doi: 10.1117/1.JBO.19.2.025004
- Darvin, M. E., Sterry, W., Lademann, J., and Vergou, T. (2011). The role of carotenoids in human skin. *Molecules* 16, 10491–10506. doi: 10.3390/molecules161210491
- DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L. M., and Little, A. C. (2010). The health of a nation predicts their mate preferences: cross-cultural variation in women's preferences for masculinized male faces. *Proc. Royal Soc. Lond.* 277, 2405–2410. doi: 10.1098/rspb.2009.2184
- Edwards, E., and Duntley, S. (1939). The pigments and color of living human skin. *Am. J. Anat.* 65, 1–33.

- Epstein, J. (2011). World domination by box office cinema admissions. *Greenash*. Retrieved August 25, 2011, from <http://greenash.net.au/thoughts/2011/07/world-domination-by-box-office-cinema-admissions/>
- Han, C., Wang, H., Hahn, A. C., Fisher, C. I., Kandrik, M., Fasolt, V., et al. (2018). Cultural differences in preferences for facial coloration. *Evol. Hum. Behav.* 39, 154–159. doi: 10.1016/j.evolhumbehav.2017.11.005
- Hancock, K. J., and Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *Br. J. Clin. Psychol.* 99, 45–56. doi: 10.1348/000712607X199981
- Hughes, D. (1999). Effects of carotenoids on human immune function. *Proc. Nutr. Soc.* 44, 713–718.
- Hughes, D. A. (2001). Dietary carotenoids and human immune function. *Nutrition* 17, 823–827. doi: 10.1016/S0899-9007(01)00638-4
- Jablonski, N. G. (2004). The evolution of human skin and skin color. *Annu. Rev. Anthropol.* 33, 585–623. doi: 10.1146/annurev.anthro.33.070203.143955
- Jablonski, N. G., and Chaplin, G. (2000). The evolution of human skin coloration. *J. Hum. Evol.* 39, 57–106. doi: 10.1006/jhev.2000.0403
- Jones, A. L. (2018). The influence of shape and colour cue classes on facial health perception. *Evol. Hum. Behav.* 39, 19–29. doi: 10.1016/j.evolhumbehav.2017.09.005
- Kourosh, A. S., Harrington, C. R., and Adinoff, B. (2010). Tanning as a behavioral addiction. *Am. J. Drug Alcohol Abuse* 36, 284–290. doi: 10.3109/00952990.2010.491883
- Krinsky, N. I., and Johnson, E. J. (2005). Carotenoid actions and their relation to health and disease. *Mol. Asp. Med.* 26, 459–516. doi: 10.1016/j.mam.2005.10.001
- Lefevre, C. E., Ewbank, M. P., Calder, A. J., von dem Hagen, E., Perrett, D. I., Hagen, E., et al. (2013). It is all in the face: carotenoid skin coloration loses attractiveness outside the face. *Biol. Lett.* 9:20130633. doi: 10.1098/rsbl.2013.0633
- Lefevre, C. E., and Perrett, D. I. (2015). Fruit over sunbed: carotenoid skin colouration is found more attractive than melanin colouration. *Q. J. Exp. Psychol.* 68, 284–293. doi: 10.1080/17470218.2014.944194
- Little, A. C., Jones, B. C., and DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. Royal Soc. B* 366, 1638–1659. doi: 10.1098/rstb.2010.0404
- Mak, A. K. Y. (2007). Advertising whiteness: an assessment of skin color preferences among urban Chinese. *Vis. Commun. Q.* 14, 144–157. doi: 10.1080/15551390701670768
- Martinkauppi, B. (2002). *Face color under varying illumination - analysis and applications*. (Oulu, Finland: Oulu University Press).
- Paiva, S. A. R., and Russell, R. M. (1999). Antioxidants and their clinical applications: beta-carotene and other carotenoids as antioxidants. *J. Am. Coll. Nutr.* 18, 426–433. doi: 10.1080/07315724.1999.10718880
- Pezdir, K., Hutchesson, M., Whitehead, R., Ozakinci, G., Perrett, D., and Collins, C. E. (2014). Can dietary intake influence perception of and measured appearance? A systematic review: dietary intake and appearance. *Nutr. Res.* 35, 175–197. doi: 10.1016/j.nutres.2014.12.002
- Pezdir, K., Rollo, M. E., Whitehead, R., Hutchesson, M. J., Ozakinci, G., Perrett, D., et al. (2017). Perceptions of carotenoid and melanin colouration in faces among young Australian adults. *Aust. J. Psychol.* 70, 85–90. doi: 10.1111/ajpy.12163
- Rao, A. V., and Rao, L. G. (2007). Carotenoids and human health. *Pharmacol. Res.* 55, 207–216. doi: 10.1016/j.phrs.2007.01.012
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208
- Samimi, D. (2005). Carotenoids: nature's cure to macular degeneration? *Nutrition Bytes* 10, 1–5. Available from: <https://escholarship.org/uc/item/5137j570>
- Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., and De Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychol. Sci.* 16, 440–444. doi: 10.1111/j.0956-7976.2005.01554.x
- Stahl, W., and Sies, H. (2003). Antioxidant activity of carotenoids. *Mol. Asp. Med.* 24, 345–351. doi: 10.1016/S0098-2997(03)00030-X
- Stephen, I. D., Coetzee, V., and Perrett, D. I. D. I. (2011). Carotenoid and melanin pigment coloration affect perceived human health. *Evol. Hum. Behav.* 32, 216–227. doi: 10.1016/j.evolhumbehav.2010.09.003
- Stephen, I. D., Coetzee, V., Smith, M. L., and Perrett, D. I. (2009a). Skin blood perfusion and oxygenation colour affect perceived human health. *PLoS One* 4:e5083. doi: 10.1371/journal.pone.0005083
- Stephen, I. D., Hiew, V., Coetzee, V., Tiddeman, B. P., and Perrett, D. I. (2017). Facial shape analysis identifies valid cues to aspects of physiological health in Caucasian, Asian, and African populations. *Front. Psychol.* 8:1883. doi: 10.3389/fpsyg.2017.01883
- Stephen, I. D., Law Smith, M. J., Stirrat, M. R., and Perrett, D. I. (2009b). Facial skin coloration affects perceived health of human faces. *Int. J. Primatol.* 30, 845–857. doi: 10.1007/s10764-009-9380-z
- Stephen, I. D., Scott, I. M. L., Coetzee, V., Pound, N., Perrett, D. I., and Penton-Voak, I. S. (2012). Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol. Hum. Behav.* 33, 260–267. doi: 10.1016/j.evolhumbehav.2011.10.003
- Stephen, I. D., and Tan, K. W. (2015). "Healthy body, healthy face? Evolutionary approaches to health perception" in *Studies in contemporary psychology: A collection of critical essays*. eds. E. Sheppard and S. Haque (Bern, Switzerland: Peter Lang Publishing).
- Swami, V., and Tovée, M. J. (2005). Female physical attractiveness in Britain and Malaysia: a cross-cultural study. *Body Image* 2, 115–128. doi: 10.1016/j.bodyim.2005.02.002
- Tan, K. W., Graf, B. A., Mitra, S. R., and Stephen, I. D. (2015). Daily consumption of a fruit and vegetable smoothie alters facial skin color. *PLoS One* 10:e0133445. doi: 10.1371/journal.pone.0133445
- Tan, K. W., Graf, B. A., Mitra, S. R., and Stephen, I. D. (2017). Impact of fresh fruit smoothie consumption on apparent health of Asian faces. *Evol. Hum. Behav.* 38, 522–529. doi: 10.1016/j.evolhumbehav.2017.02.004
- Tan, C. B. Y., Stephen, I. D., Whitehead, R., and Sheppard, E. (2012). You look familiar: how Malaysian Chinese recognize faces. *PLoS One* 7:e29714. doi: 10.1371/journal.pone.0029714
- Tan, K. W., Tiddeman, B., and Stephen, I. D. (2018). Skin texture and colour predict perceived health in Asian faces. *Evol. Hum. Behav.* 39, 320–335. doi: 10.1016/j.evolhumbehav.2018.02.003
- Tapiero, H., Townsend, D. M., and Tew, K. D. (2004). The role of carotenoids in the prevention of human pathologies. *Biomed. Pharmacother.* 58, 100–110. doi: 10.1016/j.biopha.2003.12.006
- Tiddeman, B., Burt, D. M., and Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *Comput. Graph. Appl.* 21, 42–50. doi: 10.1109/38.946630
- Tovée, M., Swami, V., Furnham, A., Mangalparsad, R., Tovée, M. J., Swami, V., et al. (2006). Changing perceptions of attractiveness as observers are exposed to a different culture. *Evol. Hum. Behav.* 27, 443–456. doi: 10.1016/j.evolhumbehav.2006.05.004
- Tovée, M. J., Furnham, A., Swami, V. (2007). "Healthy body equals beautiful body? Changing perceptions of health and attractiveness with shifting socioeconomic status" in *The Body Beautiful*. eds. V. Swami and A. Furnham (London: Palgrave Macmillan).
- Whitehead, R. D., Coetzee, V., Ozakinci, G., and Perrett, D. I. (2012a). Cross-cultural effects of fruit and vegetable consumption on skin color. *Am. J. Public Health* 102, 212–213. doi: 10.2105/AJPH.2011.300495
- Whitehead, R. D., Ozakinci, G., and Perrett, D. I. (2012b). Attractive skin coloration: harnessing sexual selection to improve diet and health. *Evol. Psychol.* 10, 842–854. doi: 10.1177/147470491201000507
- Whitehead, R. D., Re, D., Xiao, D., Ozakinci, G., and Perrett, D. I. (2012c). You are what you eat: within-subject increases in fruit and vegetable consumption confer beneficial skin-color changes. *PLoS One* 7:e32988. doi: 10.1371/journal.pone.0032988
- Wyszecki, G., and Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae*. 2nd ed. New York: John Wiley & Sons.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Tan and Stephen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vocal Parameters of Speech and Singing Covary and Are Related to Vocal Attractiveness, Body Measures, and Sociosexuality: A Cross-Cultural Study

Jaroslava Varella Valentova^{1*}, Petr Tureček², Marco Antonio Corrêa Varella¹, Pavel Šebesta³, Francisco Dyonisio C. Mendes⁴, Kamila Janaina Pereira¹, Lydie Kubicová³, Petra Stolařová³ and Jan Havlíček²

¹ Department of Experimental Psychology, Institute of Psychology, University of São Paulo, São Paulo, Brazil, ² Faculty of Science, Charles University, Prague, Czechia, ³ Faculty of Humanities, Charles University, Prague, Czechia, ⁴ Department of Basic Psychological Processes, University of Brasília, Brasília, Brazil

OPEN ACCESS

Edited by:

Ian Stephen,
Macquarie University, Australia

Reviewed by:

Justin Kyle Mogilski,
University of South Carolina
Salkehatchie, United States
Katarzyna Pisanski,
University of Sussex, United Kingdom

*Correspondence:

Jaroslava Varella Valentova
jaroslava@usp.br

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 31 October 2018

Accepted: 20 August 2019

Published: 22 October 2019

Citation:

Valentova JV, Tureček P,
Varella MAC, Šebesta P,
Mendes FDC, Pereira KJ, Kubicová L,
Stolařová P and Havlíček J (2019)
Vocal Parameters of Speech
and Singing Covary and Are Related
to Vocal Attractiveness, Body
Measures, and Sociosexuality:
A Cross-Cultural Study.
Front. Psychol. 10:2029.
doi: 10.3389/fpsyg.2019.02029

Perceived vocal attractiveness and measured sex-dimorphic vocal parameters are both associated with underlying individual qualities. Research tends to focus on speech but singing is another highly evolved communication system that has distinct and universal features with analogs in other species, and it is relevant in mating. Both speaking and singing voice provides relevant information about its producer. We tested whether speech and singing function as “backup signals” that indicate similar underlying qualities. Using a sample of 81 men and 86 women from Brazil and the Czech Republic, we investigated vocal attractiveness rated from speech and singing and its association with fundamental frequency (F0), apparent vocal tract length (VTL), body characteristics, and sociosexuality. F0, VTL, and rated attractiveness of singing and speaking voice strongly correlated within the same individual. Lower-pitched speech in men, higher-pitched speech and singing in women, individuals who like to sing more, and singing of individuals with a higher pitch modulation were perceived as more attractive. In men, physical size positively predicted speech and singing attractiveness. Male speech but not singing attractiveness was associated with higher sociosexuality. Lower-pitched male speech was related to higher sociosexuality, while lower-pitched male singing was linked to lower sociosexuality. Similarly, shorter speech VTL and longer singing VTL predicted higher sociosexuality in women. Different vocal displays function as “backup signals” cueing to attractiveness and body size, but their relation to sexual strategies in men and women differs. Both singing and speech may indicate evolutionarily relevant individual qualities shaped by sexual selection.

Keywords: human voice, song, vocal attractiveness, fundamental frequency, sociosexuality, fitness indicators, music, voice modulation

INTRODUCTION

Speech and singing are among the most common vocal productions in adult humans and their presence seems to be universally shared across modern human populations (Brown, 1991). It is assumed that they have a common ancestor (Brown, 2001, 2017; Mithen, 2005) which evolved into two specialized systems of structured vocal communication (Lehmann et al., 2009). It also seems

that prosody, the musical part of speech which conveys mainly emotional information, is rooted already in the origins of both spoken and sung vocal production (Filippi, 2016; Brown, 2017). It has recently been shown that speech and singing may have diverged from a protolanguage and split in two systems based on their communicative function. In particular, when referential and emotional functions are introduced into an artificial communication system, the system diverges into speech- and music-like vocalizations, respectively (Ma et al., 2019). Moreover, despite a vast variability across cultures, the function of specific kinds of songs (e.g., a love song) is cross-culturally comprehensible based on their structural form (Mehr et al., 2018). Interestingly, both human and bird songs tend to employ similar descending/arched melodic contour despite substantial differences in absolute pitch and duration, which indicates similar underlying motor constraints across cultures and species (Savage et al., 2017).

Singing and speech differ in the use of vocal anatomy (Sundberg, 1977, 2018), require different patterns of breathing (Leanderson et al., 1987), and neuroanatomy of production and appreciation is likewise specific to each of the two domains (Zatorre and Baum, 2012). Cognitive processing of speech and singing is also specific for each domain, as shown in patients with amusia who have intact speech processing and patients with aphasia who have no impairment of musical capacities (Peretz and Coltheart, 2003). Despite the different design features, such as the arbitrariness of speaking and regular beat and discrete set of pitches in singing, the two domains share some further features, such as hierarchical structure and complexity (Fitch, 2006). Moreover, both speaking and singing voice provide relevant information about the producer's gender, identity, location, emotional state, and behavioral tendencies (Weninger et al., 2011) and individuals can identify others based on their speech and singing (Trehub et al., 2009).

While spoken language is mostly specific to humans and language-like forms of vocalization exist in a few other animals (prairie dogs, dolphins, etc.) (Slobodchikoff et al., 1991; Janik, 2013), singing has its parallels in many other species. The capacity for learning complex songs, new sequences and sounds has arisen independently in birds (songbirds, hummingbirds, and parrots) and mammals (whales, seals, and humans) (Fitch, 2005). Since Darwin's (1871) groundbreaking works, sexual selection has been viewed as one of the most important factors that drove the evolution of singing as a way of attracting the opposite sex and advertising individual qualities. There is a large body of research showing the importance of singing in mating success across various avian and mammalian species (e.g., Searcy and Andersson, 1986). In some species, singing seems to function as an honest signal of underlying individual qualities, so that e.g., lower-pitched songs advertise a larger body size (Hall et al., 2013). In humans, irrespective of their original adaptive value, speaking and singing can likewise be considered honest signals that meet the four requisite criteria (Smith and Bird, 2000). They both require a long time for maturation, practice, and learning (Welch, 2006), their production is energetically costly because they rapidly fade (Fitch, 2006), they can suffer from noise interference, and require intense breathing (Leanderson et al.,

1987). Both speech and singing are easily perceptible by most people, are used in mating-relevant contexts, such as courtship (White et al., 2018), can increase individual mating success, and both can serve as cues to genetic qualities of the producer (Miller, 2000). There are also some significant differences between the two: singing requires higher vocal control (Zarate, 2013) and is more demanding than speech because singers need to tailor the subglottal pressure to both pitch and loudness (Sundberg, 1977, 2003). Singing can also be louder than speech, involving more muscle activity (Åkerlund and Gramming, 1994; Leanderson et al., 1987), and it includes a performative context (Fitch, 2006) which attracts more attention and is thus socially riskier. People even tend to abbreviate their singing performance in front of supposedly expert audience (Garland and Brown, 1972). It is thus well possible that singing is even harder to fake as an honest signal of underlying individual qualities than speech is, thus serving as an ornament that can affect the quantity or quality of sexual partners.

Human voice plays an important role in mate preferences and intrasexual competition (Puts, 2010; Pisanski and Feinberg, 2019), but so far, most research on human voice attractiveness and its indicators focused on speech. Some vocal parameters, especially the fundamental frequency (F0), differ between males and females of many species, with humans exhibiting an even greater sexual dimorphism than other primates (Puts et al., 2016). F0 is produced by vibrations of the vocal folds within the larynx and together with the corresponding harmonics is perceived as voice pitch (Pisanski et al., 2016). On average, men produce lower-pitched voices than women: this is due to the effects of testosterone during puberty which thickens and lengthens male vocal folds and thereby lowers the F0 (Pisanski and Feinberg, 2019). From a more general perspective, vocal sexual dimorphism is supposed to be at least in part the result of intrasexual competition, especially in the context of male-male competition (e.g., Puts, 2010). Indeed, men with lower-pitched voices are perceived as older, taller, heavier, more masculine, and more dominant than men with higher-pitched voices (Collins, 2000; Feinberg et al., 2005; Puts et al., 2006, 2007; Pereira et al., 2019). And similarly, women with lower-pitched voices are perceived as more dominant (Borkowska and Pawlowski, 2011), and both men and women with lower-pitched voices reported higher leadership capacities (Klofstad et al., 2012).

Aside from intrasexual competition, intersexual selection may have also played a role in shaping sex differences in voice. There is robust evidence that women prefer relatively low-pitched male speaking voices, while men prefer relatively high-pitched female voices (for a review, see Pisanski and Feinberg, 2019). Nevertheless, the relationship between male and female F0 and attractiveness is non-linear: the most attractive male voices are around 96 Hz and the most attractive female ones up to 280 Hz (Borkowska and Pawlowski, 2011; Saxton et al., 2015). Importantly, preferences for lower- and higher-pitched voices in men and women, respectively can be specific to certain contexts and individuals, such as short-term relationships (Little et al., 2002), coupled women (Valentová et al., 2013), and nulliparous women (Apicella and Feinberg, 2009), and in some populations that can even be inverted (Shirazi et al., 2018). Moreover, recent

evidence suggests that lower-pitched female voices are perceived as attractive (Babel et al., 2014), and women actively lower their voices when speaking to attractive men or when willing to sound attractive (Hughes et al., 2014; Pisanski et al., 2018; but see Fraccaro et al., 2011). Lower pitched voices in women can thus signal their immediate interest and/or sexual appetite.

In line with the fitness indicator hypothesis within the sexual selection theory, vocal characteristics can convey information about the underlying qualities of voice producers, e.g., information about their health and reproductive potential. For example, men with relatively low-pitched voices exhibit low cortisol and high testosterone levels, which are related to immunoreactivity (Evans et al., 2008; Hodges-Simeon et al., 2015; Puts et al., 2016). Moreover, among North American men, a lower-pitched voice is associated with more female sexual partners (Puts, 2005), and lower-pitched male Hadza hunter-gatherers have on average a higher number of offspring (Apicella et al., 2007). Furthermore, both men and women with more attractive voices reported more sexual partners, extra-pair copulations, and earlier age of the first sex (Hughes et al., 2004), which are all considered proxies of potentially higher reproductive success.

Moreover, voice attractiveness is associated with several body measures that develop under the influence of sex-specific hormones and are thus viewed as indicators of genetic and developmental quality, and subsequently also the reproductive fitness of the individual. For example, voice attractiveness is positively associated with the shoulders-to-hip ratio in men and negatively associated with the waist-to-hip ratio in women (Hughes et al., 2004). Low pitched male voices are linked to larger body size, especially weight and height, to a particular body shape (shoulder and chest circumference, shoulder-to-hip ratio) (Evans et al., 2006), and arm strength (Puts et al., 2011). Nevertheless, a recent meta-analysis had shown that compared to other vocal parameters, voice pitch is not a reliable predictor of height in adults of the same sex (Pisanski et al., 2014) and it is a poor predictor of body weight, shape, and strength (Collins, 2000; Collins and Missing, 2003; Bruckert et al., 2006; Evans et al., 2006; Sell et al., 2010; Vukovic et al., 2010; Pisanski et al., 2016; Raine et al., 2019).

Formants, on the other hand, which are the resonant frequencies of the vocal tract, are more constrained by the anatomical structures related to body size. Formants are anatomically and functionally dissociated from fundamental frequency and are therefore a more reliable indicator of body size and shape both in humans and in numerous other mammalian species (Pisanski et al., 2014). Formants are also sexually dimorphic, whereby men show lower formant frequencies than women (Pisanski et al., 2016). Individuals who produce lower formant frequencies are perceived as more physically dominant (Puts et al., 2007) and women who produce higher formant dispersion are perceived as flirtatious and attractive by both men and women (Puts et al., 2011). Individual vocal characteristics thus may provide cues to different bodily traits and sexual behaviors linked to individual's potential reproductive success.

Importantly, voice is a dynamic behavioral display which can be both intentionally and involuntarily modulated under

specific situations so as to express or exaggerate ecologically relevant traits, including emotions (Pisanski et al., 2016). For example, both men and women change their voice when speaking to infants (Foulkes et al., 2005; Broesch and Bryant, 2015) and this specific infant-directed speech affects attention and communicative outcomes of the children (Rowe, 2012; Spinelli et al., 2017). Similarly, women modulate voice pitch when speaking to attractive men (Fraccaro et al., 2011; Hughes et al., 2014; Pisanski et al., 2018) and voices of both men and women who speak to an attractive individual are perceived as more attractive by others (Leongómez et al., 2014). Also, people can volitionally increase their vocal tract length (as estimated from formant frequencies) and decrease fundamental frequency to imitate a larger body size, and vice versa (Pisanski et al., 2016). The overall prosody of speech can be effectively modulated when expressing different emotions, such as high, loud, and fast prosody while feeling happy, and the opposite pattern while being sad (for review, see Brown, 2017). Interestingly, the same vocal modulation appears when expressing emotions by music, which suggests that both displays may convey similar information (Juslin and Laukka, 2003; Zatorre and Baum, 2012).

Although both singing production and perception is a scientific research field in its own right (Sundberg, 2003), singing accuracy is related to several loci on chromosome 4 and exhibits 40% heritability (Park et al., 2012), and singing frequently features in mating contexts (e.g., as serenades and love songs, see Dukes et al., 2003; Levitin, 2008), it tends to be overlooked by psychological research on voice attractiveness. As an exception, one study found that women who were judged as good singers based solely on the audio recordings were also independently rated as more attractive based on soundless video recordings (Wapnick et al., 1997). This is in line with research which shows that in women, attractiveness and masculinity-femininity ratings based on different modalities are correlated (e.g., Valentova et al., 2017c; Pereira et al., 2019). Nevertheless, further research is needed to test to what extent are the perceptual characteristics of speech and singing voice intercorrelated and whether both vocal displays function as backup signals, i.e., as signals that indicate similar underlying qualities, rather than multiple messages, i.e., signals that indicate different qualities of individuals (see Johnstone, 1996; Bro-Jørgensen, 2010). To the best of our knowledge, only one study tested the attractiveness of speech and singing in women and it concluded that attractiveness rated from both vocal displays is correlated and in both cases increases with voice pitch (Isenstein, 2016). This can be viewed as indicating that different vocal displays may serve as backup signals.

Aims of the Current Study

In the current study, we tested whether certain perceptual singing and speaking characteristics (perceived attractiveness, voice pitch, and formant frequencies) serve as cues to specific individual physical and behavioral qualities. Since singing production is more costly than speech, one could predict that the perceived attractiveness of singing would be a stronger indicator of individual quality than the attractiveness of speech. We have therefore tested the association between the attractiveness of singing and speech and selected body fitness indicators (body

size and shape). We have also tested the relation between attractiveness ratings of both vocal displays and sociosexuality, which we used as a proxy of a short-term sexual strategy that may, especially in men, lead to increased reproductive success. We have further investigated how selected vocal parameters (voice pitch and vocal tract length as estimated from formant frequencies) mediate the possible associations between the vocal attractiveness, body cues, and sociosexuality.

Further, we tested whether the capacity to modulate the voice and singing experience may influence the rated vocal attractiveness. We hypothesized that both singing experience and a higher ability to modulate voice would lead to a more attractive vocal production.

Additionally, we tested for possible differences in vocal parameters between the sexes in two distinct populations, a Brazilian and a Czech one. So far, very little cross-cultural research has been conducted on evolutionarily relevant aspects of voice characteristics and perceptions. Majority of that research was conducted in the United States, Western and Central Europe (for review, see Pisanski and Feinberg, 2019). Studies comparing more populations with different physical, cultural, and linguistic compositions are thus needed to increase generalization of results. For example, although most North American and European studies concluded that women prefer lower-pitched male voices, Filipino women seem to follow the opposite pattern (Shirazi et al., 2018). In our study, we employed two sets of participants using sampling in one South American and one Central European population (Brazil and Czech, respectively), which differ widely as to their history, culture, ethnicity, and demographic data, and which both also differ from Western European and North American societies. Moreover, these populations also differ in several body measures, such as height and weight (e.g., Varella et al., 2014; Valentova et al., 2016), facial and body hair in men (Valentova et al., 2017b), while self-rated breast size, buttock size, and WHR in women is the same in both (Valentova et al., 2017a). Furthermore, Brazilian population reports a significantly higher sociosexuality than the Czech population (Varella et al., 2014). Both populations are also linguistically different: Brazilian Portuguese is a Latin language while Czech belongs to Slavic languages. Previous studies reported that several vocal parameters differ between the different linguistic groups (Mennen et al., 2012). The two populations thus offer an interesting opportunity to analyze vocal production and perception and its relation to body measures and sociosexuality.

METHODS

Target Participants

The final sample was composed of 40 Brazilian men ($M = 23.70$ years; $SD = 3.67$, range 19–34) and 44 women ($M = 23.91$ years; $SD = 4.99$, range 18–35) recruited at the University of São Paulo, in São Paulo city, and 33 Czech men ($M = 22.45$ years; $SD = 2.35$, range 18–28) and 35 women ($M = 22.37$ years; $SD = 2.57$, range 19–29), recruited at the Charles University, Prague. We selected predominantly heterosexual participants (0–2 on a Kinsey scale) because

individuals with different sexual orientations can show variation in several vocal parameters (Kachel et al., 2018) which can be detected even by naïve listeners (Valentova and Havlíček, 2013).

Procedure

In both countries, each participant consented to take part in a broader study (see, Varella et al., 2014; Valentova et al., 2017c). Participants completed questionnaires, we took body measurements, standardized facial and body photographs, and recorded videos of both speech and singing. Only data relevant for this study are described below. Brazilians are not allowed to receive financial reward but Czech participants received remuneration amounting to 300 CZK (approximately 13 USD). The project was approved by the Charles University IRB (2011/07).

Questionnaires

Participants completed a sociodemographic questionnaire and the Revised Sociosexual Orientation Inventory (SOI-R; Penke and Asendorpf, 2008). The SOI-R measures an individual's willingness to engage in uncommitted sex. It consists of nine items (e.g., "With how many different partners did you have sexual intercourse on one and only one occasion?"), loading into three subscales of sociosexual behavior, attitudes, and desire. They also answered, on a 10-point scale, how much they liked to sing (1 = not at all, 10 = very much). We used this information as a motivational factor that may influence singing frequency, singing training, and thus singing experience, as shown in Busch (2013).

Vocal Recordings

Vocal samples were recorded under standardized conditions, in a closed and quiet room, and all by one researcher. For all recordings, we used a professional digital stereo Olympus LS-100 Multi-Track Linear PCM recorder, whereby the participants' lips were approximately 10 cm from the microphone. When performing the vocal tasks, all participants were seated on a chair. First, participants were informed about the whole recording procedure: this information was printed for them. After a small vocal exercise to warm-up the voice and get used to being recorded, participants read a short sentence using standardized names across all participants. In Brazil, all men and women, respectively, pronounced "*Oi, meu nome é Pedro/Ana, e eu sou de Belo Horizonte*," while Czech men and women, respectively, said "*Jmenuji se Petr/Petra a pocházím z Havlíčkova Brodu*" (Hi, my name is Petr/Pedro/Petra/Ana and I come from Belo Horizonte/Havlíčkův Brod). Subsequently, they sang the first part of "Happy Birthday" (in the Brazilian Portuguese version "*Parabéns para você, nesta data querida, muitas felicidades, muitos anos de vida*," in the Czech version "*Hodně štěstí zdraví, hodně štěstí zdraví, hodně štěstí, milý Honzo, hodně štěstí zdraví*"). Finally, they first read and then sang the first stanza of their national anthem (the verbal content of speech and singing was thus matched).

To minimize raters' overload, we extracted parts of the national anthem using SoundForge 8.0 software. In the Brazilian sample, we extracted the first two lines of the national anthem ("*Ouviram do Ipiranga as margens plácidas, de um povo heróico o brado retumbante*"), while for the Czech participants,

we extracted the third and fourth line, which unlike the first two lines are not repetition of each other (“*Voda hučí po lučinách, bory šumí po skalinách*”). Only these recordings were subsequently rated by independent participants and analyzed for vocal parameters. All participants spoke their native language, i.e., either Brazilian Portuguese or Czech. None of the participants reported any serious vocal or respiratory problem at the time of the data collection.

Happy Birthday was selected because it is cross culturally known and commonly sung in intimate and emotionally loaded social situations, usually with the family, friends, and romantic partners, and it has been used in research on singing previously (e.g., Christiner and Reiterer, 2013). The national anthem is also widely known within each country, it is relatively unconnected to mating context and is thus more neutral.

Recordings were analyzed using Praat software (Boersma and Weenink, 2013) for mean, minimal, and maximal fundamental frequency (F0), and the first four formants (F1–F4). F0 is the rate of vocal folds vibration perceived as an overall voice pitch. We used an autocorrelation algorithm with parameters set to a pitch floor of 75 Hz and pitch ceiling of 300 Hz for men, and a pitch floor of 100 Hz and pitch ceiling of 500 Hz for women, because these are the appropriate boundaries for analyzing adult voices recommended by the software developers (Boersma and Weenink, 2013). All other values were set to default. Average speech F0 per recording ranged between 92.47 (Corresponding to musical note F#₂, here F note is heightened by semitone, which is indicated by #) and 177.70 Hz (F₃) in men, and between 164.10 (E₃) and 253.10 Hz (B₃) in women. For singing, F0 ranged between 103.60 (G#₂) and 208.50 Hz (G#₃) in men, and between 168.5 (E₃) and 348.20 Hz (F₄) in women. All F0 were transformed to perceptual pitch expressed in a semitone difference between A₄ (440 Hz) and F0 using a standard formula $12\log_2(F0/440)$. This scale is based on standard music notation and reflects the logarithmic nature of human pitch perception, where both A₃ (–12, 220 Hz) and A₅ (12, 880 Hz) are at an equal octave distance (12 semitones) from A₄. We subtracted the minimal F0 from the maximal F0 of each recording to obtain its perceptual range in semitones. Average speech range per recording ranged between 4.61 and 21.07 semitones in men and between 5.34 and 27.61 semitones in women, while the singing range ranged between 6.76 and 23.74 semitones in men, and between 8.76 and 27.84 semitones in women. F0 and ranges were averaged for each participant across recordings for speech and singing separately.

Apparent vocal tract length (VLT) was calculated from the first four formants (F1–F4) according to a formula described in Pisanski et al. (2014). F1 to F4 were measured in Praat using semiautomated approach. First, recordings were preprocessed by Vocal Toolkit’s “Extract voiced and unvoiced” script (Corrette, 2019) and subsequently only the voiced parts were used for further formants analysis. Second, formants were analyzed by Burg method with recommended preset values and maximum formant levels of 5000 and 5500 Hz for men and women, respectively. In each recording from the list of results were omitted readings suggesting presence of silence and erroneous readings. F1 to F4 levels are represented by median of remaining formants readings.

Subsequently, formant spacing (ΔF) was estimated as a slope of the linear regression line with an intercept set to 0 from a relationship

$$F_i = \frac{(2i - 1)}{2} \Delta F$$

where “*i*” refers to the formant number. Apparent vocal tract length was derived from formant spacing using

$$VTL(\Delta F) = \frac{c}{2\Delta F}$$

where $c = 33.500$ cm/s is the speed of sound in a uniform tube with one end closed.

Anthropometry

We measured participants’ body height in centimeters, weight in kilograms, and body characteristics previously found to be associated with vocal attractiveness, namely the circumference of the shoulders, waist, and hips (Dixon et al., 2003; Stulp et al., 2013; Valentova et al., 2014, 2016, 2017a). Then we computed the waist-to-shoulder ratio (WSR) in men and waist-to-hip ratio (WHR) in women (for details on the procedure, see Varella et al., 2014).

Vocal Ratings

An independent sample of heterosexual raters anonymously judged voice attractiveness of all vocal recordings of individuals of the opposite sex on a 7-point scale (1 = not at all attractive, 7 = very attractive) using Rater software (facelab.org). All raters reported being predominantly heterosexual (0–2 on a Kinsey scale). Brazilian raters (51 men: $M = 22$ years, $SD = 3.4$ years; 59 women: $M = 22.1$ years, $SD = 3.4$) were recruited among the students of the University of Brasília, while the Czech raters (46 men: $M = 21.7$ years, $SD = 1.9$; 47 women: $M = 20.6$ years, $SD = 1.1$) were recruited at the Charles University, Prague. The rating took place in an empty classroom, each voice recording containing the relevant phrase was presented only once using headphones and with unmanipulated volume. Each rater evaluated either all Brazilian or all Czech recordings. For instance, one Brazilian rater rated all Czech recordings, while another Brazilian rater rated all Brazilian recordings. The recordings were divided into eight blocks (two speech and two singing recordings, Brazilian and Czech sample) and randomized within each block. Interrater agreement (Cronbach’s α) was high in all recording \times rater set combinations (min $\alpha = 0.79$) (For a full overview of Cronbach’s α , see **Supplementary Material**). Pearson correlations between average attractiveness ratings of Czech and Brazilian raters were high for both speech [$r = 0.694$, 95%CI (0.602,0.768) $p < 0.001$] and singing [$r = 0.788$, 95%CI (0.719,0.841) $p < 0.001$]. We have therefore used as a unit of analysis the mean rating of attractiveness for each target across all raters.

Statistical Analyses

All analyses were conducted using R 3.5.1 software, and SPSS version 21 (IBM Corp., Armonk, NY, United States). To

explore associations between the measured and rated voice parameters in speech and song, we ran parametric correlations (Pearson correlation) and paired *t*-tests to test for possible differences between the two vocal displays.

Relationships between the four exogenous variables (waist-to-hip or waist-to-shoulders ratio, height, weight, and age), mediating acoustic qualities (speech and singing F0 and range), speech and singing attractiveness, and the total sociosexuality score were investigated using path analysis. The structural model contained 6 correlations and 38 regression coefficients. Analysis was conducted using *sem()* function from the *lavaan* package. Because of small parameters/observations ratio (as low as 1.66 in the male sample), robust *p* values were obtained using Monte Carlo simulation. The distribution of expected correlation/regression coefficients was derived from 10,000 simulation runs, where the full model was estimated on a randomized dataset. The issue of influential points was avoided by jackknife resampling. Removing one observation at a time, we extracted sets of all measures including standardized model estimates and *p* values. Coefficients which remained significant regardless of the removed data points are emphasized in the main article, while full results are reported in the **Supplementary Material**. Path invariance was tested from the χ^2 difference between configural invariant, where structure is restricted to be equal between the groups, and path invariant, where all coefficients are restricted to be equal between the groups, with degrees of freedom corresponding to the number of estimated parameters. Path invariance was evaluated between men and women and subsequently between Czech and Brazilian participants within each sex. Interrater agreement was evaluated using Cronbach's α calculated using *alpha()* function from the *psych* package (the code is available at <https://github.com/costlysignalling/Speech-and-singing-attractiveness>).

Further, to test for the possible effect of voice experience on rated voice attractiveness, we assessed non-parametric correlations (Kendall rank correlation indicated by coefficient τ) between the rated attractiveness of both spoken and sung recordings and how much the participants liked to sing. To test the voice modulation hypothesis, we computed the absolute difference between singing and speaking F0, singing and speaking F0 range, and the absolute difference between singing and speaking VTL, which gave us an index of (dis)similarity of these vocal parameters between the two vocal displays. The higher the absolute difference, the larger the difference between speech and singing, and thus the higher vocal modulation. We further correlated these absolute differences with attractiveness ratings, separately for men and women. In these analyses, we did not control for multiple comparisons across tests, because the samples were independent.

Additionally, we used General Linear Models (GLM) to test for possible effects of sex, age, and country on voice attractiveness ratings. Similarly, to test whether mean F0, range F0, and VTL of speech and singing differ between men and women or between Brazilian and Czech participants, we performed a multivariate GLM with mean F0 and F0 range as dependent variables and sex and country of targets as factors. Due to a limited samples size, we

evaluated only simple models. The effect size displayed is a partial Eta-squared (η_p^2).

RESULTS

The Effect of Targets' Sex and Country on Spoken and Sang F0, F0 Range, and VTL

We found large effects of targets' sex on all vocal parameters; mean speech F0 ($F = 1074.30$, $df = 1, 153$, $p < 0.001$, $\eta_p^2 = 0.878$), mean speech F0 range ($F = 14.12$, $df = 1, 153$, $p < 0.001$, $\eta_p^2 = 0.086$), VTL as measured from speech ($F = 2114.02$, $df = 1, 153$, $p < 0.001$, $\eta_p^2 = 0.934$), mean singing F0 ($F = 736.84$, $df = 1, 153$, $p < 0.001$, $\eta_p^2 = 0.831$), mean singing F0 range ($F = 7.00$, $df = 1, 153$, $p = 0.009$, $\eta_p^2 = 0.045$), and VTL as measured from singing ($F = 1537.91$, $df = 1, 153$, $p < 0.001$, $\eta_p^2 = 0.911$). Estimated marginal means revealed that women had a higher F0 and F0 range and shorter VTL than men (for mean values, see **Table 1**). There was also a significant effect of the target country on speech F0 range ($F = 4.31$, $df = 1, 153$, $p = 0.040$, $\eta_p^2 = 0.028$), VTL as measured from speech ($F = 10.49$, $df = 1, 153$, $p = 0.001$, $\eta_p^2 = 0.065$), and VTL as measured from singing ($F = 6.59$, $df = 1, 153$, $p = 0.011$, $\eta_p^2 = 0.042$). Estimated marginal means show that Czech participants had a lower speech F0 range and longer VTL than the Brazilian participants (see **Table 1** for details).

It is worth noting that the average VTL measures for men and women (**Table 1**) compare to population-level averages (Pisanski et al., 2014).

Comparisons Between Speaking and Singing Voice

F0 measured from speech was strongly positively correlated with F0 measured from singing in both men ($r = 0.800$, $N = 73$, $p < 0.001$) and women ($r = 0.607$, $N = 79$, $p < 0.001$). F0 range measured from speech was correlated with F0 range measured from singing in men ($r = 0.408$, $N = 73$, $p < 0.001$) but not in women ($r = 0.160$, $N = 79$, $p < 0.159$). Vocal tract length (VTL) as estimated from formant frequencies was strongly positively correlated between speech and singing in both men ($r = 0.808$, $N = 81$, $p < 0.001$) and women ($r = 0.764$, $N = 85$, $p < 0.001$). Vocal attractiveness rated from speech and singing was also strongly positively correlated in both men ($r = 0.720$, $N = 73$, $p < 0.001$) and women ($r = 0.674$, $N = 79$, $p < 0.001$). Paired *t*-test revealed that voices rated from speech were judged significantly higher on attractiveness than voices rated from singing in both men ($t = 6.66$, $df = 72$, $p < 0.001$) and women ($t = 3.85$, $df = 78$, $p \leq 0.001$).

Structural Models

The model which analyzes the fundamental frequency is not path-invariant with respect to the sex of individuals ($\chi^2 = 117.03$, $df = 44$, $p < 0.001$) but is path-invariant with respect to participants' nationality ($\chi^2 = 49.58$, $df = 44$, $p = 0.26$ in men, $\chi^2 = 60.68$, $df = 44$, $p = 0.05$ in women). Results are therefore

TABLE 1 | Mean fundamental frequency (F0) and the range of fundamental frequency (F0 range) in semitones, and VTL (in centimeters) in men and women.

	Men			Women		
	Brazilian (N = 42)	Czech (N = 35)	Total (N = 77)	Brazilian (N = 45)	Czech (N = 36)	Total (N = 81)
Mean F0 – speech (SD)	–22.15 (2.13)	–22.63 (1.84)	–22.37 (2.01)	–13.19 (1.28)	–13.00 (1.53)	–13.11 (1.39)
Mean F0 – singing (SD)	–19.71 (2.49)	–20.51 (2.26)	–20.07 (2.40)	–10.50 (2.05)	–10.07 (2.08)	–10.31 (2.06)
Mean F0 range – speech (SD)	11.98 (2.52)	11.02 (2.90)	11.55 (2.72)	14.25 (4.23)	12.97 (3.55)	13.69 (3.97)
Mean F0 range – singing (SD)	14.65 (2.52)	14.24 (2.42)	14.47 (2.46)	15.70 (3.25)	15.74 (2.96)	15.72 (3.10)
VTL – speech (SD)	17.32 (0.49)	17.57 (0.51)	17.44 (0.51)	14.18 (0.30)	14.38 (0.39)	14.27 (0.35)
VTL – singing (SD)	16.80 (0.53)	17.18 (0.64)	16.98 (0.61)	13.93 (0.38)	13.95 (0.34)	13.94 (0.36)

reported separately for men and women but jointly for Czech and Brazilian participants.

Using path analysis (see **Supplementary Tables S6, S7** for full models), we found that in men, lower-pitched speech was rated as more attractive (**Figure 1**). The same held of singing, but this relationship did not reach statistical significance. In men, a broader speech range, but not singing range, was rated as more attractive. Attractive speech was positively associated with the total SOI, but this relationship failed to maintain its stability in jackknife resampling. The total SOI was directly connected to a lower F0 in speech and higher F0 in singing. Body weight had a strong and positive direct effect on perceived speech and singing attractiveness. Age had a negative effect on speech attractiveness but the effect failed to remain stable under jackknifing (see **Supplementary Table S8**).

Higher-pitched female voices (both in speech and singing) were rated as more attractive. No other relationship except for correlation between height and weight was significant (see **Supplementary Tables S7, S9**).

The additional model that analyzed vocal tract length (VTL) was not path-invariant with respect to the sex of individuals ($\chi^2 = 109.44$, $df = 44$, $p < 0.001$) but was path-invariant with respect to participants' nationality at least in women ($\chi^2 = 66.99$, $df = 44$, $p = 0.01$ in men, $\chi^2 = 59.18$, $df = 44$, $p = 0.06$ in women). Results are reported separately for men and women but jointly for Czech and Brazilian participants for a better comparison with the original model that employs the F0.

Many relationships in the structural model remained similar when we replaced average F0 with apparent VTL (**Figure 2**). Nevertheless, the VTL failed to predict speech or singing attractiveness reliably. In women, we observed a reverse relationship between speech and singing VTL and the total SOI. In this model, however, these relationships were stronger because the potentially mediating path between VTL and attractiveness was weaker. This was possibly due to the fact that in the first model, which relied on average fundamental frequency together with the F0 range, both measurements of vocal quality were based on the same characteristic (F0 – either as average or as a difference between minimum and maximum), which in effect allowed us to partition out their respective contributions to speech and singing attractiveness better. The model with VTL, which tightly correlated with average F0, lowered the partial correlations beyond the threshold of statistical significance. All the relationships were, however, in the direction that would be

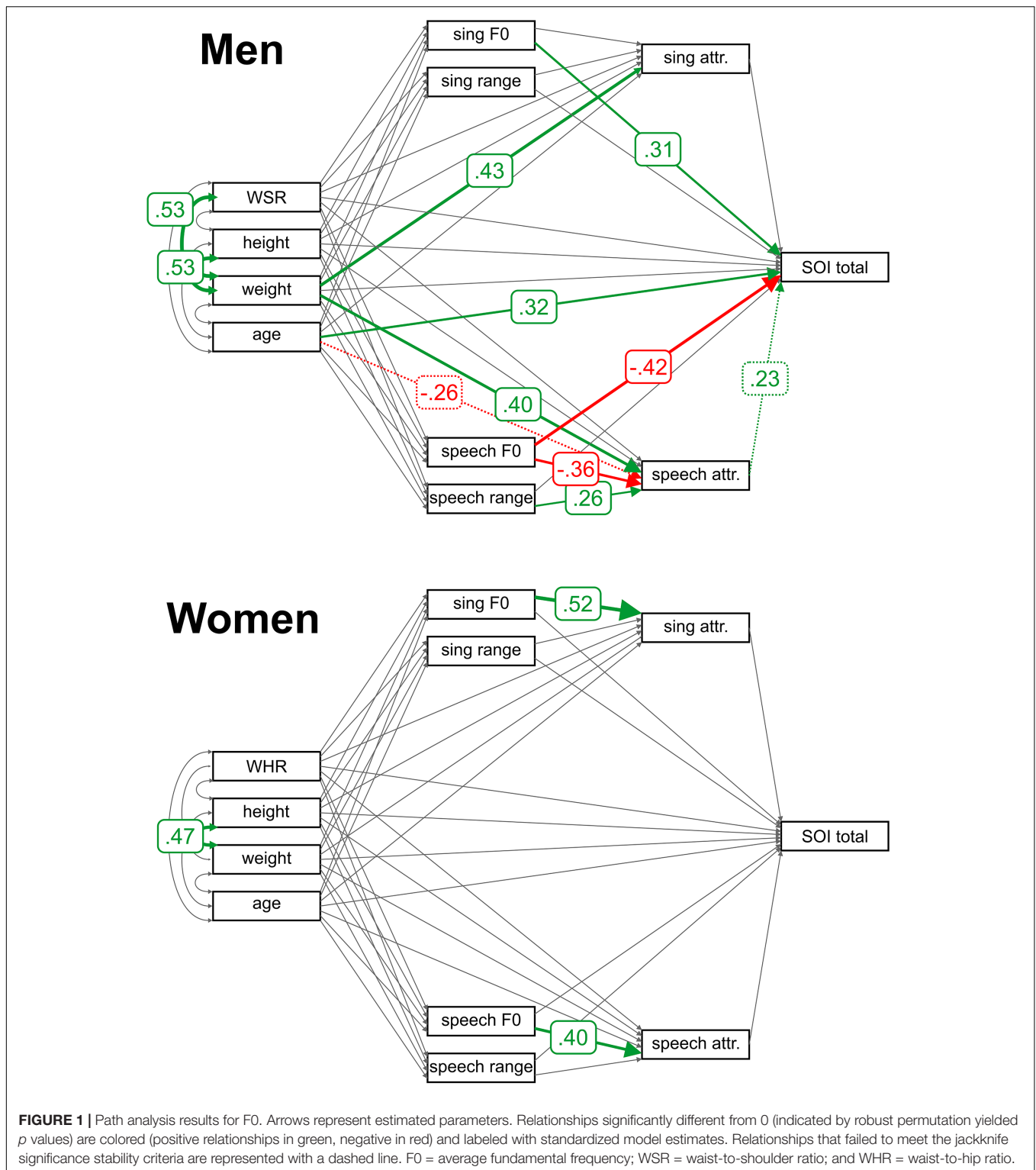
expected based on the strong negative correlation between VTL and mean F0 (See **Supplementary Tables S10–S12**).

The Effect of Singing Experience and Voice Modulation on Voice Attractiveness

Non-parametric correlations showed a positive association between how much men liked to sing and attractiveness as rated from both speech ($\tau = 0.253$, $N = 87$, $p < 0.001$) and singing ($\tau = 0.277$, $N = 87$, $p < 0.001$). In women, this association was rather weak and significant only in singing attractiveness ($\tau = 0.171$, $N = 90$, $p = 0.024$) but not in speech attractiveness ($\tau = 0.101$, $N = 91$, $p = 0.183$). Furthermore, the absolute difference of F0 between speech and singing was positively correlated with how much men and women liked to sing ($\tau = 0.255$, $N = 90$, $p = 0.001$; $\tau = 0.281$, $N = 93$, $p < 0.001$, respectively). Moreover, the absolute difference of F0 was positively associated with rated singing attractiveness in both men ($\tau = 0.177$, $N = 87$, $p = 0.015$) and women ($\tau = 0.294$, $N = 90$, $p < 0.001$) but not significantly associated with speech attractiveness in either men ($\tau = 0.123$, $N = 87$, $p = 0.092$) or women ($\tau = 0.118$, $N = 90$, $p = 0.101$). Finally, the absolute difference of F0 was weakly positively associated with sociosexuality in men ($\tau = 0.139$, $N = 80$, $p = 0.069$) but not in women ($\tau = 0.036$, $N = 84$, $p = 0.632$). There were no significant correlations with the absolute difference between spoken and sung F0 range or VTL, rated attractiveness, and sociosexuality.

The Effect of Targets' Sex and Country on Voice Attractiveness Ratings From Speech and Singing

Test of between-subjects effects of the GLM model showed significant main effect of sex of targets on attractiveness rated both from speech ($F = 13.84$, $df = 1$, 157 , $p < 0.001$, $\eta_p^2 = 0.082$) and singing ($F = 36.48$, $df = 1$, 157 , $p < 0.001$, $\eta_p^2 = 0.192$). Estimated marginal means revealed that the voices of female participants were rated as more attractive based on both speech (mean rating = 3.89, $SD = 0.65$) and singing (mean rating = 3.82, $SD = 0.73$) than the voices of male participants (mean ratings = 3.48, $SD = 0.66$; and 3.11, $SD = 0.72$, respectively). There was no effect of country.



DISCUSSION

Using a cross-cultural sample of men and women, we have shown that speech and singing attractiveness are strongly correlated. We also found a strong correlation between the fundamental

frequency (F0), F0 range, and vocal tract length (VTL) in both vocal displays. In men, low-pitched speech was rated as attractive and a similar trend was observed in singing. Furthermore, both vocal displays were invariably associated with body size (but not shape) and differently associated with sociosexuality.

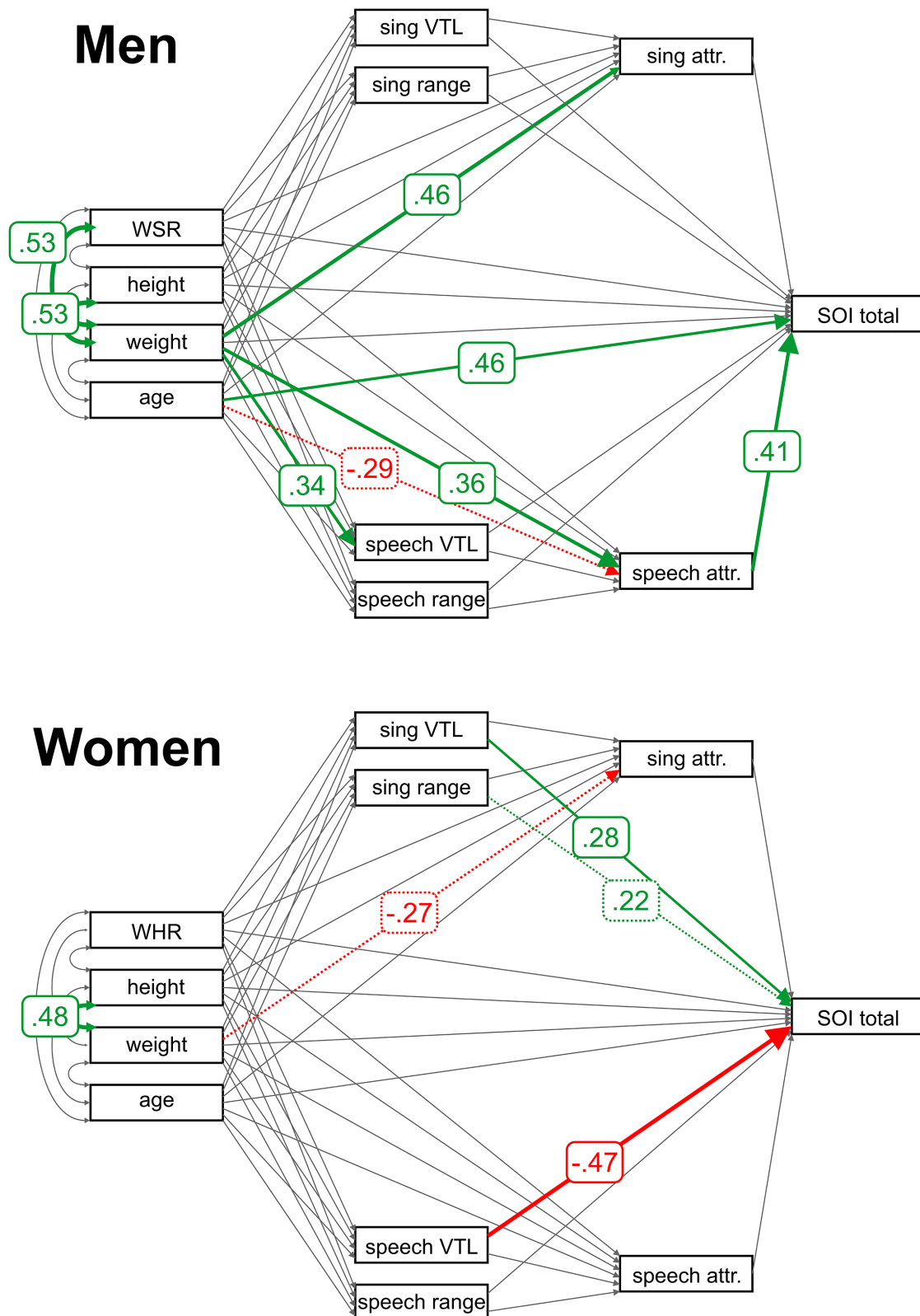


FIGURE 2 | Path analysis results for VTL. Arrows represent estimated parameters. Relationships significantly different from 0 (indicated by robust permutation yielded p values) are colored (positive relationships in green, negative in red) and labeled with standardized model estimates. Relationships that failed to fulfill the jackknife significance stability criteria are represented with a dashed line. VTL = apparent vocal tract length; WSR = waist-to-shoulder ratio; and WHR = waist-to-hip ratio.

In women, both high-pitched singing and speaking voice predicted vocal attractiveness, and similarly to men, VTL as measured from singing and speech was differently associated with sociosexuality. Most results were invariant with respect to participants' nationality, which indicates a degree of universality.

Our results partly support the hypothesis that speech and singing work as backup signals. They share many vocal parameters, such as fundamental frequency, its range and formant frequencies, which could lead to similar attractiveness ratings in both vocal displays (for similar results, see Isenstein, 2016). Both studied vocal displays thus covary in their production and perception and can transmit similar information to listeners. This is in line with previous studies which show that women's cross-modal attractiveness or masculinity as rated from faces and spoken voices are intercorrelated, although no such correlation was found in men (Valentova et al., 2016; Pereira et al., 2019).

Nevertheless, we also found some features which are specific to the singing and speaking voice. For example, male speech attractiveness, but not singing attractiveness, is associated with higher sociosexuality (for similar results, see Hughes et al., 2004). The observed absence of association between singing attractiveness and male sociosexuality may suggest that singing voice is not part of the repertoire of short-term sexual strategy, at least in the two studied populations, which does not, however, exclude the possibility that it may be used to foster long-term relationships. Further, in line with previous studies, lower F0 in speech was directly connected to higher sociosexuality in men (e.g., Puts, 2005), while lower F0 in singing was connected to lower sociosexuality. Again, this could point to possible use of singing vocal display rather for committed long-term sexual strategy, which needs to be tested in future studies.

Further, although a high F0 in both speech and singing predicted vocal attractiveness in women, only low speech F0 was rated as attractive in men, although a similar non-significant trend appeared also in singing. This is in line with a study that found no difference in the attractiveness ratings of high- and low-pitched performances of famous singers (Neumann et al., 2008). Nevertheless, when analyzing the relative vocal parameters (difference in voice pitch between spoken and sung voice of the same person), we found that the singing voice of individuals who are capable of a higher pitch modulation is perceived as more attractive. In accordance with the handicap theory, individuals who can produce a larger difference between their spoken baseline and singing performance can thus benefit in terms of higher attractiveness and consequently potentially higher fitness. In line with this, men who modulated their voice pitch more had a tendency for higher sociosexuality, and men who like to sing more had more attractive voices. Both singing experience and higher capacity of voice modulation are thus linked to male attractiveness and sexuality.

Interestingly, in our study speech was on average rated as more attractive than singing. This can indicate that the standards for evaluation are higher in the singing domain, whereby singing abilities (e.g., singing in-tune), which are 40% heritable (Park et al., 2012), and were not tested in this study, may have influenced this difference. Nevertheless, another study found higher attractiveness ratings of singing than in speech in women

and found no association between attractiveness ratings and singing quality (Isenstein, 2016). More studies are clearly needed to discern and determine the overall pattern.

We found that body weight was a strong positive predictor of both speech and singing attractiveness in men and a weak negative predictor of singing attractiveness in women (for similar results, see e.g., Sell et al., 2010; Xu et al., 2013; Šebesta et al., 2017). Weight also positively predicted VTL as estimated from speech in men, which is likewise in line with previous studies (for a review, see Pisanski et al., 2014). Some studies found differences in several vocal parameters (F0, voice pressure, perceptual voice quality) as a function of body weight, whereby heavier individuals have lower-pitched voices of more attractive perceptual quality (Barsties et al., 2013; Jost et al., 2018). The link between decrease in F0 and increase in body weight could be driven by hormonal factors, since for example in men, increased amount of fat tissue relates to lower testosterone levels (Zumoff et al., 1990; Tchernof et al., 1995). On the other hand, body weight may be due to not only body fat but also muscularity, which are both correlated with body size. Since the male body is composed relatively more by muscles than by fat tissue, one could speculate that vocal attractiveness provides a reliable cue specifically to muscularity, but future studies should assess the contribution of individual body components to vocal attractiveness. We also predicted a stronger association between body size and singing attractiveness but our results did not confirm this hypothesis. In humans, as in some songbirds (Hall et al., 2013), different vocal manifestations can thus serve as a cue to body size but not to body shape. This is in line with the finding that lower-pitched voice affects the perception of physical dominance (Puts et al., 2007).

Although women report that they like to sing more than men (Varella et al., 2010), and women and men both prefer sexual partners who demonstrate some music abilities (Kaufman et al., 2016), we found no association between singing or speaking voice attractiveness and sociosexuality or body indicators in women. This is contrary to previous studies (e.g., Hughes et al., 2004) which reported that women with attractive speaking voices had a lower waist-to-hip ratio, age of first sex, and a higher total number of sexual partners. Nevertheless, we found that shorter VTL measured from speech and longer VTL measured from singing predicted higher sociosexuality in women (for similar results in men, see Hodges-Simeon et al., 2011). This is comparable to our finding obtained for men when we analyzed the fundamental frequency. Generally speaking, individuals with sex-typical speech parameters and sex atypical singing parameters have higher sexual success (see, Bártová et al., 2019, for similar results on higher sociosexuality and gender non-conformity), which further supports the handicap hypothesis. Interestingly, there was no effect of the VTL on voice attractiveness and no effect of voice attractiveness on sociosexuality in women. Women's tendency for sexual variety thus does not seem to be defined by how attractive they appear to the opposite sex. Access to sexual partners in individuals who display honest signals can be influenced by other mechanisms, such as intra-sexual competition (Varella et al., 2017; Ostrander et al., 2018).

This is the first study whose aim was to test the potential involvement of intersexual selection on different vocal

displays on a cross-cultural sample of men and women (for intrasexual selection, see Raine et al., 2018; Šebesta et al., 2019). Although we used four different vocal recordings (standardized self-presentation, singing of “Happy Birthday,” and reading and singing of the national anthem), they do not represent the full range of human speech or singing. Standardized songs, such as “Happy Birthday,” are likely to limit pitch dynamics and range and thereby obscure or dampen the individual differences in pitch and voice modulation which might otherwise provide important cues to fitness.

Studies using different vocal recordings, such as spontaneous speech and singing, singing of more mating-relevant songs, or wordless singing, should be undertaken. This might be why some of our predictions were not supported. It is for instance possible that a link between quality indicators and singing attractiveness becomes apparent in more demanding singing that involves complex rhythms, melody, or range (Charlton, 2014). The production of such demanding songs could be viewed as costly signaling and therefore serve as a more reliable indicator than the relatively undemanding songs employed in this study. Moreover, future studies should also perform more fine-tuned vocal analyses to compare both singing and speech (Šebesta et al., 2019).

It also ought to be taken into account that our samples in both countries were recruited from middle-class university student populations in the largest cities of both countries. They were thus not representative of the local populations and moreover compared only two countries. More cross-cultural comparisons are needed to test the generalizability potential of our current findings (see, Moshontz et al., 2018 for multi-lab psychological studies). Finally, as correlations between Czech and Brazilian raters were high, we pooled the ratings together, and did not analyze potential in-group and out-group effects, which might be addressed in future studies.

To conclude, we expected that singing would be a stronger indicator of individual body characteristics and sexuality than speech but our results show that cross-culturally, speech and singing seem to work rather in concert, i.e., as backup signals. Attractiveness of both singing and speaking voice is perceived in a similar way and is connected to a higher pitch in women and a lower pitch in men. Moreover, in men, speaking and singing both serve as similar cues to body indicators. On the other hand, the relation between speaking and singing voice and sociosexuality works in opposite ways in both men and women. Developmental pathways leading to sex-typical or atypical speaking and singing voice and sexuality should be addressed in future studies. In general, singing, together with other vocalizations, should be taken into account in evolutionary literature on voice production and perception.

REFERENCES

- Åkerlund, L., and Gramming, P. (1994). Average loudness level, mean fundamental frequency, and subglottal pressure: comparison between female singers and nonsingers. *J. Voice* 8, 263–270. doi: 10.1016/s0892-1997(05)80298-x
- Apicella, C. L., and Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter – gatherers. *Proc. R. Soc. B* 276, 1077–1082. doi: 10.1098/rspb.2008.1542

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Charles University IRB with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Charles University IRB (2011/07).

AUTHOR CONTRIBUTIONS

JV and JH developed the study concept and MV expanded it. JV, MV, FM, KP, LK, and PS collected the data. JV performed the analysis of F0 and F0 range of the vocal stimuli. PŠ performed the formant analyses during revisions of the manuscript. JV and PT performed the data analysis and interpretation jointly with MV and JH. JV and MV drafted the manuscript. PT and JH provided the critical revisions. JV, MV, JH, PŠ, and PT worked on the revised version of the manuscript. All authors approved the final version of the manuscript for submission.

FUNDING

JH was supported by the Charles University Research Centre program UNCE 204056. MV was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), number PNPd 33002010037P0 – MEC/CAPES.

ACKNOWLEDGMENTS

We are indebted to all volunteers for their participation and Anna Pilátová, Ph.D. for English proofreading. We are grateful for Tiago Leal Dutra de Andrade for helping with collecting data during the ratings phase in Brasília. We further thank Prof. Dr. Vera S. R. Bussab for enabling the initial data collection phase at the University of São Paulo. We also thank the reviewers who offered valuable and critical suggestions of improvements.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02029/full#supplementary-material>

- Apicella, C. L., Feinberg, D. R., and Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett.* 3, 682–684. doi: 10.1098/rsbl.2007.0410
- Babel, M., McGuire, G., and King, J. (2014). Towards a more nuanced view of vocal attractiveness. *PLoS One* 9:e88616. doi: 10.1371/journal.pone.0088616
- Barsties, B., Verfaillie, R., Roy, N., and Maryn, Y. (2013). Do body mass index and fat volume influence vocal quality, phonatory range, and aerodynamics in females? *Codas* 25, 310–318.

- Bártová, K., Štěrbová, Z., Varella, M. A. C., and Valentova, J. V. (2019). Femininity in men and masculinity in women is positively related to sociosexuality. *Pers. Individ. Differ.* 152:109575. doi: 10.1016/j.paid.2019.109575
- Boersma, P., and Weenink, D. (2013). *Praat: Doing Phonetics by Computer [Computer Program]*, Version 5.3.42. Available at: <http://www.fon.hum.uva.nl/praat/> doi: 10.1590/s2317-17822013000400003 (accessed July, 2019).
- Borkowska, B., and Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Anim. Behav.* 82, 55–59. doi: 10.1016/j.anbehav.2011.03.024
- Broesch, T. L., and Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *J. Cogn. Dev.* 16, 31–43. doi: 10.1080/15248372.2013.833923
- Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: animal communication in a world in flux. *Trends Ecol. Evol.* 25, 292–300. doi: 10.1016/j.tree.2009.11.003
- Brown, D. E. (1991). *Human Universals*. New York, NY: McGraw Hill.
- Brown, S. (2001). “The ‘musilanguage’ model of music evolution,” in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown, (Cambridge, MA: MIT Press), 271–300.
- Brown, S. (2017). A joint prosodic origin of language and music. *Front. Psychol.* 8:1894. doi: 10.3389/fpsyg.2017.01894
- Bruckert, L., Liénard, J.-S., Lacroix, A., Kreutzer, M., and Leboucher, G. (2006). Women use voice parameters to assess men’s characteristics. *Proc. R. Soc. B Biol. Sci.* 273, 83–89. doi: 10.1098/rspb.2005.3265
- Busch, S. L. (2013). *Beyond Singer vs. Non-singer in Singing, Health and Well-being: Development and Testing of the Singing Experience Scale*. Doctoral dissertation, Carleton University, Ottawa, ON.
- Charlton, B. D. (2014). Menstrual cycle phase alters women’s sexual preferences for composers of more complex music. *Proc. R. Soc. B* 281:20140403. doi: 10.1098/rspb.2014.0403
- Christner, M., and Reiterer, S. M. (2013). Song and speech: examining the link between singing talent and speech imitation ability. *Front. Psychol.* 4:874. doi: 10.3389/fpsyg.2013.00874
- Collins, S. A. (2000). Men’s voices and women’s choices. *Anim. Behav.* 60, 773–780.
- Collins, S. A., and Missing, C. (2003). Vocal and visual attractiveness are related in women. *Anim. Behav.* 65, 997–1004. doi: 10.1006/anbe.2003.2123
- Corrette, R. (2019). *Praat Vocal Toolkit*. Available at: <http://www.praatvocaltoolkit.com> (accessed July, 2019).
- Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London: J. Murray.
- Dixon, A. F., Halliwell, G., East, R., Wignarajah, P., and Anderson, M. J. (2003). Masculine somatotype and hirsuteness as determinants of sexual attractiveness to women. *Arch. Sex. Behav.* 32, 29–39. doi: 10.1023/a:1021889228469
- Dukes, R. L., Bisel, T. M., Borega, K. N., Lobato, E. A., and Owens, M. D. (2003). Expressions of love, sex, and hurt in popular songs: a content analysis of all-time greatest hits. *Soc. Sci. J.* 40, 643–650. doi: 10.1016/s0362-3319(03)00075-2
- Evans, S., Neave, N., and Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males: an evolutionary explanation for a deep male voice. *Biol. Psychol.* 72, 160–163. doi: 10.1016/j.biopsycho.2005.09.003
- Evans, S., Neave, N., Wakelin, D., and Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiol. Behav.* 93, 783–788. doi: 10.1016/j.physbeh.2007.11.033
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim. Behav.* 69, 561–568. doi: 10.1016/j.anbehav.2004.06.012
- Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: a comparative approach to the emergence of language. *Front. Psychol.* 7:1393. doi: 10.3389/fpsyg.2016.01393
- Fitch, W. T. (2005). The evolution of language: a comparative review. *Biol. Philos.* 20, 193–203. doi: 10.1007/s10539-005-5597-1
- Fitch, W. T. (2006). The biology and evolution of music: a comparative perspective. *Cognition* 100, 173–215. doi: 10.1016/j.cognition.2005.11.009
- Foulkes, P., Docherty, G. J., and Watt, D. (2005). Phonological variation in child-directed speech. *Language* 81, 177–206. doi: 10.3109/02699206.2015.1115555
- Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., et al. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *J. Evol. Psychol.* 9, 57–67. doi: 10.1556/jep.9.2011.33.1
- Garland, H., and Brown, B. R. (1972). Face-saving as affected by subjects’ sex, audiences’ sex and audience expertise. *Sociometry* 35, 280–289.
- Hall, M. L., Kingma, S. A., and Peters, A. (2013). Male songbird indicates body size with low-pitched advertising songs. *PLoS One* 8:e56717. doi: 10.1371/journal.pone.0056717
- Hodges-Simeon, C. R., Gaulin, S. J., and Puts, D. A. (2011). Voice correlates of mating success in men: examining “contests” versus “mate choice” modes of sexual selection. *Arch. Sex. Behav.* 40, 551–557. doi: 10.1007/s10508-010-9625-0
- Hodges-Simeon, C. R., Gurven, M., and Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. *Evol. Hum. Behav.* 36, 294–302. doi: 10.1016/j.evolhumbehav.2015.01.002
- Hughes, S. M., Dispenza, F., and Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evol. Hum. Behav.* 25, 295–304. doi: 10.1016/j.evolhumbehav.2004.06.001
- Hughes, S. M., Mogilski, J. K., and Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *J. Nonverbal. Behav.* 38, 107–127. doi: 10.1007/s10919-013-0163-z
- Isenstein, S. (2016). *Singing Voice Attractiveness*. Master thesis, McMaster University, Hamilton, ON.
- Janik, V. M. (2013). Cognitive skills in bottlenose dolphin communication. *Trends Cogn. Sci.* 17, 157–159. doi: 10.1016/j.tics.2013.02.005
- Johnstone, R. A. (1996). Multiple displays in animal communication: ‘backup signals’ and ‘multiple messages’. *Philos. Trans. R. Soc. B* 351, 329–338. doi: 10.1098/rstb.1996.0026
- Jost, L., Fuchs, M., Loeffler, M., Thiery, J., Kratzsch, J., Berger, T., et al. (2018). Associations of sex hormones and anthropometry with the speaking voice profile in the adult general population. *J. Voice* 32, 261–272. doi: 10.1016/j.jvoice.2017.06.011
- Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814. doi: 10.1037/0033-2909.129.5.770
- Kachel, S., Radtke, A., Skuk, V. G., Zäske, R., Simpson, A. P., and Steffens, M. C. (2018). Investigating the common set of acoustic parameters in sexual orientation groups: a voice averaging approach. *PLoS One* 13:e0208686. doi: 10.1371/journal.pone.0208686
- Kaufman, S. B., Kozbelt, A., Silvia, P., Kaufman, J. C., Ramesh, S., and Feist, G. J. (2016). Who finds Bill Gates sexy? Creative mate preferences as a function of cognitive ability, personality, and creative achievement. *J. Creat. Behav.* 50, 294–307. doi: 10.1002/jocb.78
- Klofstad, C. A., Anderson, R. C., and Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proc. R. Soc. B Biol. Sci.* 279, 2698–2704. doi: 10.1098/rspb.2012.0311
- Leanderson, R., Sundberg, J., and Von Euler, C. (1987). Breathing muscle activity and subglottal pressure dynamics in singing and speech. *J. Voice* 1, 258–261. doi: 10.1016/s0892-1997(87)80009-7
- Lehmann, C., Welker, L., and Schiefenhövel, W. (2009). Towards an ethology of song: a categorization of musical behaviour. *Music. Sci.* 13, 321–338. doi: 10.1177/1029864909013002141
- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., et al. (2014). Vocal modulation during courtship increases proceptivity even in naive listeners. *Evol. Hum. Behav.* 35, 489–496. doi: 10.1016/j.evolhumbehav.2014.06.008
- Levitin, D. J. (2008). *The World in Six Songs: How the Musical Brain Created Human Nature*. New York, NY: Penguin.
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., and Perrett, D. I. (2002). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proc. R. Soc. B Biol. Sci.* 269, 1095–1100. doi: 10.1098/rspb.2002.1984
- Ma, W., Fiveash, A., and Thompson, W. F. (2019). Spontaneous emergence of language-like and music-like vocalizations from an artificial protolanguage. *Semiotica* 229, 1–23. doi: 10.1515/sem-2018-0139 [Epub ahead of print].
- Mehr, S. A., Singh, M., York, H., Glowacki, L., and Krasnow, M. M. (2018). Form and function in human song. *Curr. Biol.* 28, 356–368. doi: 10.1016/j.cub.2017.12.042

- Mennen, I., Schaeffler, F., and Docherty, G. (2012). Cross-language differences in fundamental frequency range: a comparison of English and German. *J. Acoust. Soc. Am.* 131, 2249–2260. doi: 10.1121/1.3681950
- Miller, G. (2000). "Evolution of human music through sexual selection," in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown, (Cambridge MA: MIT Press), 329–360.
- Mithen, S. (2005). *The Singing Neanderthal*. London: Weidenfeld & Nicholson.
- Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., et al. (2018). The psychological science accelerator: advancing psychology through a distributed collaborative network. *Adv. Meth. Pract. Psychol. Sci.* 1, 501–515.
- Neumann, K., Schunda, P., Lehmann, C., and Euler, H. A. (2008). Attractiveness of the high male speaking and singing voice. *Paper Presented at Conference Choice for the Voice of the British Voice Association*, London.
- Ostrander, G. M., Pipitone, R. N., and Shoup-Knox, M. L. (2018). Interactions between observer and stimuli fertility status: endocrine and perceptual responses to intrasexual vocal fertility cues. *Horm. Behav.* 98, 191–197. doi: 10.1016/j.yhbeh.2017.12.008
- Park, H., Lee, S., Kim, H. J., Ju, Y. S., Shin, J. Y., Hong, D., et al. (2012). Comprehensive genomic analyses associate UGT8 variants with musical ability in a Mongolian population. *J. Med. Genet.* 49, 747–752. doi: 10.1136/jmedgenet-2012-101209
- Penke, L., and Asendorpf, J. B. (2008). Beyond global sociosexual orientations: a more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *J. Pers. Soc. Psychol.* 95, 1113–1135. doi: 10.1037/0022-3514.95.5.1113
- Pereira, K. J., da Silva, C. S. A., Havlíček, J., Kleisner, K., Varella, M. A. C., Pavlovič, O., et al. (2019). Femininity-masculinity and attractiveness – Associations between self-ratings, third-party ratings and objective measures. *Pers. Individ. Differ.* 147, 166–171. doi: 10.1016/j.paid.2019.04.033
- Peretz, I., and Coltheart, M. (2003). Modularity of music processing. *Nat. Neurosci.* 6, 688–691. doi: 10.1038/nn1083
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J., and Reby, D. (2016). Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* 20, 304–318. doi: 10.1016/j.tics.2016.01.002
- Pisanski, K., and Feinberg, D. R. (2019). "Vocal attractiveness," in *Oxford Handbook of Voice Perception*, eds S. Frühholz, and P. Belin, (New York, NY: Oxford University Press), 607–625.
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., Röder, S., Andrews, P. W., et al. (2014). Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* 95, 89–99. doi: 10.1016/j.anbehav.2014.06.011
- Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., and Reby, D. (2018). Voice pitch modulation in human mate choice. *Proc. R. Soc. B* 285:20181634. doi: 10.1098/rspb.2018.1634
- Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* 26, 388–397. doi: 10.1016/j.evolhumbehav.2005.03.001
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Puts, D. A., Barndt, J. L., Welling, L. L., Dawood, K., and Burriss, R. P. (2011). Intrasexual competition among women: vocal femininity affects perceptions of attractiveness and flirtatiousness. *Pers. Individ. Differ.* 50, 111–115. doi: 10.1016/j.paid.2010.09.011
- Puts, D. A., Gaulin, S. J. C., and Verdolini, K. V. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* 27, 283–296. doi: 10.1016/j.evolhumbehav.2005.11.003
- Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., et al. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proc. R. Soc. B Biol. Sci.* 283:20152830. doi: 10.1098/rspb.2015.2830
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., and Gaulin, S. J. C. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* 28, 340–344. doi: 10.1016/j.evolhumbehav.2007.05.002
- Raine, J., Pisanski, K., Bond, R., Simner, J., and Reby, D. (2019). Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech. *PLoS One* 14:e0213034. doi: 10.1371/journal.pone.0213034
- Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J., and Reby, D. (2018). Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* 4, 273–280. doi: 10.1016/j.isci.2018.05.002
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Dev.* 83, 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x
- Savage, P. E., Tierney, A. T., and Patel, A. D. (2017). Global music recordings support the motor constraint hypothesis for human and avian song contour. *Music Percept.* 34, 327–334. doi: 10.1525/mp.2017.34.3.327
- Saxton, T. K., Mackey, L. L., McCarty, K., and Neave, N. (2015). A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. *Behav. Ecol.* 27, 512–519. doi: 10.1093/beheco/arv178
- Searcy, W. A., and Andersson, M. (1986). Sexual selection and the evolution of song. *Ann. Rev. Ecol. Sys.* 17, 507–533.
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., et al. (2017). Voices of Africa: acoustic predictors of human male vocal attractiveness. *Anim. Behav.* 127, 205–211. doi: 10.1016/j.anbehav.2017.03.014
- Šebesta, P., Třebický, V., Fialová, J., and Havlíček, J. (2019). Roar of a champion: loudness and voice pitch predict perceived fighting ability but not success in MMA fighters. *Front. Psychol.* 10:859. doi: 10.3389/fpsyg.2019.00859
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. B* 277, 3509–3518. doi: 10.1098/rspb.2010.0769
- Shirazi, T. N., Puts, D. A., and Escasa-Dorne, M. J. (2018). Filipino women's preferences for male voice pitch: intra-individual, life history, and hormonal predictors. *Adapt. Hum. Behav. Physiol.* 4, 188–206. doi: 10.1007/s40750-018-0087-2
- Slobodchikoff, C. N., Kiriazis, J., Fischer, C., and Creef, E. (1991). Semantic information distinguishing individual predators in the alarm calls of Gunnison's prairie dogs. *Anim. Behav.* 42, 713–719. doi: 10.1016/s0003-3472(05)80117-4
- Smith, E. A., and Bird, R. L. B. (2000). Turtle hunting and tombstone opening: public generosity as costly signaling. *Evol. Hum. Behav.* 21, 245–261. doi: 10.1016/S1090-5138(00)00031-3
- Spinelli, M., Fasolo, M., and Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Devel. Rev.* 44, 1–18. doi: 10.1016/j.dr.2016.12.001
- Stulp, G., Buunk, A. P., Pollet, T. V., Nettle, D., and Verhulst, S. (2013). Are human mating preferences with respect to height reflected in actual pairings? *PLoS One* 8:e54186. doi: 10.1371/journal.pone.0054186
- Sundberg, J. (1977). The acoustics of the singing voice. *Sci. Am.* 236, 82–91. doi: 10.1038/scientificamerican0377-82
- Sundberg, J. (2003). Research on the singing voice in retrospect. *TMH QPSR* 45, 11–22.
- Sundberg, J. (2018). "The singing voice," in *The Oxford Handbook of Voice Perception*, eds S. Frühholz, and P. Belin, (Oxford: Oxford University Press), 117–142.
- Tchernof, A., Després, J. P., Bélanger, A., Dupont, A., Prud'homme, D., Moorjani, S., et al. (1995). Reduced testosterone and adrenal C19 steroid levels in obese men. *Metabolism* 44, 513–519. doi: 10.1016/0026-0495(95)90060-8
- Trehub, S. E., Plantinga, J., and Brcic, J. (2009). Infants detect cross-modal cues to identity in speech and singing. *Ann. N. Y. Acad. Sci.* 1169, 508–511. doi: 10.1111/j.1749-6632.2009.04851.x
- Valentova, J., Roberts, S. C., and Havlíček, J. (2013). Preferences for facial and vocal masculinity in homosexual men: the role of relationship status, sexual restrictiveness, and self-perceived masculinity. *Perception* 42, 187–197. doi: 10.1068/p6909
- Valentova, J. V., Bártová, K., Štěrbová, Z., and Varella, M. A. C. (2016). Preferred and actual relative height are related to sex, sexual orientation, and dominance: evidence from Czech Republic and Brazil. *Pers. Individ. Differ.* 100, 145–150. doi: 10.1016/j.paid.2016.01.049
- Valentova, J. V., Bártová, K., Štěrbová, Z., and Varella, M. A. C. (2017a). Influence of sexual orientation, population, homogamy, and imprinting-like effect on preferences and choices for female buttock size, breast size and

- shape, and WHR. *Pers. Individ. Differ.* 104, 313–319. doi: 10.1016/j.paid.2016.08.005
- Valentova, J. V., and Havlíček, J. (2013). Perceived sexual orientation based on vocal and facial stimuli is linked to self-rated sexual orientation in Czech men. *PLoS One* 8:e82417. doi: 10.1371/journal.pone.0082417
- Valentova, J. V., Stulp, G., Třebický, V., and Havlíček, J. (2014). Preferred and actual relative height among homosexual male partners vary with preferred dominance and sex role. *PLoS One* 9:e86534. doi: 10.1371/journal.pone.0086534
- Valentova, J. V., Varella, M. A. C., Bártová, K., Štěřbová, Z., and Dixon, B. J. W. (2017b). Mate preferences and choices for facial and body hair in heterosexual women and homosexual men: influence of sex, population, homogamy, and imprinting-like effect. *Evol. Hum. Behav.* 38, 241–248. doi: 10.1016/j.evolhumbehav.2016.10.007
- Valentova, J. V., Varella, M. A. C., Havlíček, J., and Kleisner, K. (2017c). Positive association between vocal and facial attractiveness in women but not in men: a cross-cultural study. *Behav. Proc.* 135, 95–100. doi: 10.1016/j.beproc.2016.12.005
- Varella, M. A. C., Ferreira, J. H. B. P., Cosentino, L. A. M., Ottoni, E., and Bussab, V. S. R. (2010). Sex differences in aspects of musicality in a Brazilian sample: adaptative hypotheses. *Cogn. Music. Arts* 4, 10–16.
- Varella, M. A. C., Valentova, J. V., and Fernández, A. M. (2017). “Evolution of artistic and aesthetic propensities through female competitive ornamentation,” in *The Oxford Handbook of Women and Competition*, ed. M. Fisher, (New York, NY: Oxford University Press), 757–783.
- Varella, M. A. C., Valentova, J. V., Pereira, K. J., and Bussab, V. S. R. (2014). Promiscuity is related to masculine and feminine body traits in both men and women: evidence from Brazilian and Czech samples. *Behav. Process.* 109, 34–39. doi: 10.1016/j.beproc.2014.07.010
- Vukovic, J., Feinberg, D. R., DeBruine, L., Smith, F. G., and Jones, B. C. (2010). Women’s voice pitch is negatively correlated with health risk factors. *J. Evol. Psychol.* 8, 217–225. doi: 10.1556/JEP.8.2010.3.2
- Wapnick, J., Darrow, A. A., Kovacs, J., and Dalrymple, L. (1997). Effects of physical attractiveness on evaluation of vocal performance. *J. Res. Music Educ.* 45, 470–479. doi: 10.2307/3345540
- Welch, G. F. (2006). “Singing and vocal development,” in *The Child as Musician: A Handbook of Musical Development*, ed. G. McPherson, (New York, NY: Oxford University Press), 311–329.
- Weninger, F., Wöllmer, M., and Schuller, B. (2011). “Automatic assessment of singer traits in popular music: Gender, age, height and race,” in *Proceedings of the 12th International Society for Music Information Retrieval Conference*, Miami, FL.
- White, J., Lorenz, H., Perilloux, C., and Lee, A. (2018). Creative casanovas: mating strategy predicts using—but not preferring—atypical flirting tactics. *Evol. Psych. Sci.* 4, 443–455. doi: 10.1007/s40806-018-0155-7
- Xu, Y., Lee, A., Wu, W. L., Liu, X., and Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One* 8:e62397. doi: 10.1371/journal.pone.0062397
- Zarate, J. M. (2013). The neural control of singing. *Front. Hum. Neurosci.* 7:237. doi: 10.3389/fnhum.2013.00237
- Zatorre, R. J., and Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS Biol.* 10:e1001372. doi: 10.1371/journal.pbio.1001372
- Zumoff, B., Strain, G. W., Miller, L. K., Rosner, W., Senie, R., Seres, D. S., et al. (1990). Plasma free and non-sex-hormone-binding-globulin bound testosterone are decreased in obese men in proportion to their degree of obesity. *J. Clin. Endocrinol. Metab.* 71, 929–931. doi: 10.1210/jcem-71-4-929

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Valentova, Tureček, Varella, Šebesta, Mendes, Pereira, Kubicová, Stolařová and Havlíček. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership