



OPEN

Linking human male vocal parameters to perceptions, body morphology, strength and hormonal profiles in contexts of sexual selection

Christoph Schild¹, Toe Aung², Tobias L. Kordsmeyer³, Rodrigo A. Cardenas⁴, David A. Puts^{2,5}✉ & Lars Penke^{3,5}✉

Sexual selection appears to have shaped the acoustic signals of diverse species, including humans. Deep, resonant vocalizations in particular may function in attracting mates and/or intimidating same-sex competitors. Evidence for these adaptive functions in human males derives predominantly from perception studies in which vocal acoustic parameters were manipulated using specialist software. This approach affords tight experimental control but provides little ecological validity, especially when the target acoustic parameters vary naturally with other parameters. Furthermore, such experimental studies provide no information about what acoustic variables indicate about the speaker—that is, why attention to vocal cues may be favored in intrasexual and intersexual contexts. Using voice recordings with high ecological validity from 160 male speakers and biomarkers of condition, including baseline cortisol and testosterone levels, body morphology and strength, we tested a series of pre-registered hypotheses relating to both perceptions and underlying condition of the speaker. We found negative curvilinear and negative linear relationships between male fundamental frequency (f_0) and female perceptions of attractiveness and male perceptions of dominance. In addition, cortisol and testosterone negatively interacted in predicting f_0 , and strength and measures of body size negatively predicted formant frequencies (P_f). Meta-analyses of the present results and those from two previous samples confirmed that f_0 negatively predicted testosterone only among men with lower cortisol levels. This research offers empirical evidence of possible evolutionary functions for attention to men's vocal characteristics in contexts of sexual selection.

Acoustic signals comprise a fundamental component of mating competition^{1–4} and are highly sexually dimorphic in many species, including many anthropoid primates. Humans in particular exhibit strong sexual dimorphism in acoustic signals⁵, such that the distributions of male and female vocal parameters related to pitch and timbre barely overlap⁶.

From hearing the voice alone, humans can assess diverse salient social characteristics of a speaker, such biological sex, age and physical strength^{7–9}. Many of these evaluations rely on inter-individual variation in specific sets of vocal parameters, including fundamental frequency and formant frequencies^{5,10}. Fundamental frequency (f_0) is the rate of vocal fold vibration during phonation and influences perceptions of pitch. Formant frequencies are resonant frequencies determined by the length and shape of the vocal tract and influence perceptions of vocal timbre.

¹Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen, Denmark. ²Department of Anthropology and Center for Brain, Behavior and Cognition, Pennsylvania State University, University Park, PA 16802, USA. ³Department of Psychology and Leibniz ScienceCampus Primate Cognition, University of Goettingen, Gosslerstrasse 14, 37073 Göttingen, Germany. ⁴Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA. ⁵These authors jointly supervised this work: David A. Puts and Lars Penke. ✉email: dap27@psu.edu; lpenke@uni-goettingen.de

Fundamental and formant frequencies are some of the most sexually dimorphic characteristics in humans, suggesting a past influence of sexual selection¹¹. Indeed, lower male f_0 predicts greater perceptions of attractiveness, dominance and masculinity^{12–14}, as well as greater mating success^{14,15} (but see¹⁶ for a null finding) and reproductive success^{15,17} (see also^{18,19}). Likewise, male formant frequencies influence perceptions of attractiveness, dominance and masculinity^{12,13,20,21}.

Despite the abundance of evidence linking acoustic parameters to perceptions relevant in mating competition, a fundamental question remains: Why have humans evolved to attend to these parameters? Costly signaling theory (originally proposed by^{22,23}, but see²⁴) which concerns the transmission of reliable information between signalers and receivers, is a useful theoretical tool to answer this question and helps us understand the maintenance of signal honesty via receiver-independent (production costs, developmental costs, maintenance costs) and receiver-dependent costs (e.g., retaliation costs, vulnerability costs; see^{25,26} for reviews). Recently, some authors^{27,28} have pointed out weak receiver-independent costs associated with men's f_0 and concluded that men's f_0 does not signal formidability. Others^{29–31} suggest that men's f_0 is likely to be partly honest.

Although f_0 influences perceptions of physical dominance, it correlates only weakly with physical strength^{6,9,32} (see²⁹ for a meta-analysis) and body height³³. Past research also points to associations with hormonal profiles in males: f_0 decreases strongly during, and higher circulating testosterone levels predict lower f_0 in men^{11,34,35} (see²⁹ for a meta-analysis). Further, the relationship between f_0 and testosterone was found to be stronger in men with lower cortisol levels⁵, a pattern that has been associated with immunocompetence³⁶. Another study³⁷ that utilized salivary immunoglobulin-A (sIgA; a marker of mucosal immunity) as a measure of immunocompetence reported that sIgA was negatively correlated with f_0 . In a similar vein, listeners assigned higher dominance ratings, but not higher health ratings, to speakers with higher self-reported health³⁸. Overall, these studies suggest, that f_0 may be a partly honest signal of condition^{29–31}. Formants are closely tied to vocal tract length and are therefore indirect, albeit weak, correlates of body size in humans^{33,39,40}. Additionally, a recent study showed significant correlations with other somatometric measures, such as body mass index and hip circumference⁴¹. However, links between formants and physical strength are equivocal^{6,32}.

In addition to the paucity of evidence concerning the information content of male voices, there are also significant gaps in knowledge concerning how men's voices may influence social perceptions. For example, because most prior studies manipulated only one acoustic parameter at a time in experimental settings, the relative importance of different parameters in forming social judgments have not been well characterized. Prior research also has primarily investigated linear relationships (Table 1), and thus it remains largely unknown whether acoustic parameters have curvilinear effects on perceptions, which have been predicted in some cases¹¹. Vocal stimuli in most prior work are also unnaturally invariant in content and motivation, with all speakers uttering a series of vowels, counting, or speaking precisely the same, often socially irrelevant, phrase; hence, the generalizability and external validity of such results depend on whether the effects they reveal persist in natural speech¹³. Finally, only a few, mostly low-powered studies (Table 1) have simultaneously shown that these acoustic parameters are related to both perceptions of attractiveness and/or dominance on the one hand and indirect measures of mate quality and formidability on the other.

Given the fundamental gaps in knowledge outlined above, we conducted a preregistered study (preregistration: <https://osf.io/nrmpf/>) to examine (1) how vocal parameters are utilized in assessing dominance and attractiveness, and (2) why using those parameters for judgments could be adaptive insofar as they are associated with indirect measures of mate quality and/or formidability. In contrast to most studies on perceived vocal attractiveness and dominance, which have used standardized voice samples (i.e. counting, vowels or standardized passages), more natural stimuli were used to augment external validity. Importantly, we use a relatively large ($N = 160$) and rich dataset, which allows relationships between vocal parameters, baseline cortisol and testosterone levels, body morphology and strength to be tested in a single sample.

Perceptions of attractiveness and dominance. Because deep male voices may display social power²⁹, threat potential¹¹, and predict greater anticipated^{42,86,87} and actual^{42,88} sexual infidelity, there may be costs as well as benefits to mating with males with masculine voices¹¹. Further, some studies suggest that the link between mean f_0 and attractiveness is weaker and rather curvilinear: Very low-pitched voices are not seen as more attractive and sometimes even less attractive as low-pitched voices^{11,64}. In line with the context-dependent nature of costs and benefits and reports from previous literature, we therefore predicted negative linear⁵ and negative quadratic¹¹ relationships between attractiveness ratings and both mean f_0 (**H1**) and formant position (P_f) (**H2**). P_f is a measure of formant structure, calculated as the average standardized formant value for the first n (usually four) formants⁶.

Masculine voices (i.e. low f_0 and P_f) have been found to be preferred by females to a greater extent in short-term compared to long-term relationship contexts^{14,89}. This might reflect an adaptive trade-off strategy in which a mate's genetic fitness, putatively indicated by masculine traits, is granted greater value in short-term contexts, whereas his expected investment and fidelity are valued more in long-term contexts^{89,90}. Consequently, we predicted stronger relationships between short-term, compared to long-term, attractiveness ratings and both mean f_0 (**H3**) and P_f (**H4**).

It has been hypothesized that deep voices display threat potential⁶; hence, we predicted negative relationships between dominance ratings and both mean f_0 (**H5**) and P_f (**H6**). According to the source-filter theory, f_0 and P_f are theoretically distinct⁹¹. They are also only weakly correlated¹⁰ and seem to convey different information about a male speaker⁶. Accordingly, we predicted f_0 and P_f to be independent predictors of both attractiveness (**H7**) and dominance (**H8**) ratings.

| No | Studies | Rater (n) | Vocalizers (n) | Perceptions evaluated | Vocalizer's condition | Natural voices | Cuvilinear tested |
|----|--|-----------|----------------|-----------------------|-----------------------|----------------|-------------------|
| 1 | Schild et al. ⁴² | 95 | 181 | Trus | Trus | + | + |
| 2 | Collins and Missing ⁴³ | 30 | 30 | Att; Age | Size | + | |
| 3 | Puts et al. ⁵ | 1126 | 548 | Att; Dom | T; C | + | |
| 4 | Raine et al. ⁴⁴ | 150 | 61 | Size | Size | + | |
| 5 | Raine et al. ⁴⁵ | 135 | 61 | Size | Size | + | |
| 6 | Rendall et al. ⁴⁶ | 163 | 68 | Size | Size | + | |
| 7 | Rosenfield et al. ¹⁵ | 84 | 4 | Att; Pres; Dom | MS | + | |
| 8 | Šebesta et al. ⁴⁷ | 62 | 93 | Att | Size | + | |
| 9 | Šebesta et al. ⁴⁸ | 63 | 40 | Dom | Size | + | |
| 10 | Simmons et al. ⁴⁹ | 30 | 44 | Att; Mas | Semen | + | |
| 11 | Valentova et al. ⁵⁰ | 203 | 152 | Att | Size | + | |
| 12 | Armstrong et al. ²⁷ | 224 | 183 | Dom; Size | Size | + | |
| 13 | Feinberg et al. ⁵¹ | 991 | 123 | Age; Att; Fem | | + | + |
| 14 | Babel et al. ⁵² | 30 | 60 | Att | | + | |
| 15 | Gregory et al. ⁵³ | 118 | 60 | Com Qual | | + | |
| 16 | Hodges-Simeon et al. ¹³ | 330 | 111 | Att; Dom | | + | |
| 17 | Knowles et al. ⁵⁴ | 180 | 32 | Cop | | + | |
| 18 | Michalsky and Schoormann ⁵⁵ | 20 | 20 | Att; Like | | + | |
| 19 | Pisanski and Rendall ⁵⁶ | 129 | 89 | Size; Att; Mas; Fem | | + | |
| 20 | Pisanski et al. ⁵⁷ | 68 | 20 | Size; Att; Mas; Fem | | + | |
| 21 | Sorokowski et al. ⁵⁸ | 39 | 51 | Comp; Auth | | + | |
| 22 | Valentova et al. ⁵⁹ | 84 | 30 | Att; Mas | | + | |
| 23 | Hill et al. ⁶⁰ | 1349 | 471 | Att | Fac Sym | | + |
| 24 | Wolff and Puts ⁶¹ | 376 | 117 | Dom | Size; T; Agg | | + |
| 25 | Shirazi et al. ⁶² | 128 | 6 | Att | E; P | | |
| 26 | Re et al. ⁶³ | 19 | 64 | Att; Mas; Fem | | | + |
| 27 | Saxton et al. ⁶⁴ | 40 | 6 | Att; Dom | | | + |
| 28 | Apicella and Feinberg ⁶⁵ | 88 | 10 | Att | | | |
| 29 | Borkowska and Pawlowski ⁶⁶ | 473 | 58 | Att; Dom | | | |
| 30 | Bruckert et al. ⁶⁷ | 64 | 55 | Att | | | |
| 31 | Feinberg et al. ⁶⁸ | 68 | 5 | Att; Dom | | | |
| 32 | Feinberg et al. ⁶⁹ | 26 | 8 | Att; Dom | | | |
| 33 | Feinberg et al. ⁷⁰ | 1759 | 6 | Pref | | | |
| 34 | Feinberg et al. ⁷¹ | 83 | 6 | Att | | | |
| 35 | Fraccaro et al. ⁷² | 179 | 8 | Att; Dom | | | |
| 36 | Hughes et al. ⁷³ | 40 | 40 | Att | | | |
| 37 | Jones et al. ⁷⁴ | 800 | 12 | Att; Dom | | | |
| 38 | Klofstad et al. ⁷⁵ | 382 | 27 | Com; Size; Trus | | | |
| 39 | Leaderbrand et al. ⁷⁶ | 48 | 4 | Att | | | |
| 40 | O'Connor et al. ⁷⁷ | 138 | 6 | Att; Inv | | | |
| 41 | Puts et al. ⁷⁸ | 86 | 111 | Dom | | | |
| 42 | Puts et al. ²⁰ | 42 | 30 | Dom | | | |
| 43 | Puts et al. ⁷⁹ | 109 | 4 | Att; Flir | | | |
| 44 | Puts ¹⁴ | 142 | 111 | Att | | | |
| 45 | Riding et al. ⁸⁰ | 54 | 9 | Att | | | |
| 46 | Suire et al. ⁸¹ | 225 | 58 | Att | | | |
| 47 | Tigue et al. ⁸² | 165 | 15 | Int; Prow; Vote | | | |
| 48 | Vukovic et al. ⁸³ | 70 | 6 | Att; Dom; Trus | | | |
| 49 | Watkins et al. ⁸⁴ | 50 | 10 | Dom | | | |
| 50 | Xu et al. ⁸⁵ | 42 | 2 | Att; Emo | | | |

Table 1. A non-exhaustive list of studies (n = 50) on human voice perception. A list of 50 studies that relate to mating-relevant perceptions of human voice was obtained via Google Scholar search. Most studies that investigate human voice perceptions tested only on perceptions (n = 35), used manipulated voice stimuli (n = 28), and tested linear relationships (n = 44). Agg = Aggressiveness; Att = Attractiveness; C = Cortisol; Com = Competent; Com Qual = Communication Quality; Cop = Cooperativeness; Dom = Dominance; Emo = Emotions; E = Estradiol; Fac Sym = Facial Symmetry; Flir = Flirtatiousness; Fem = Femininity; Int = Integrity; Inv = Investing; Mas = Masculinity; MS = Mating Success; P = Progesterone; Pref = Preference; Pres = Prestige; Prow = Prowess; T = Testosterone; Trus = Trustworthiness; + = Presence.

Indirect measures of mate quality and formidability. Previous studies^{34,35} linked lower f_0 to higher circulating testosterone levels, and more recently this relationship was found to be stronger in men with lower cortisol levels³, a result seemingly consistent with the stress-linked immunocompetence handicap hypothesis that f_0 honestly signals a speaker's physical condition³⁶. We therefore predicted a negative relationship between mean f_0 and testosterone (**H9**) and predicted that this relationship would be attenuated by high baseline cortisol (**H10**).

Formants have been shown to relate moderately to body height, a phenotype that is relevant in both intra- and intersexual selective contexts⁹². We therefore predicted a negative relationship between P_f and body height (**H11**).

Exploratory analyses. In addition to these preregistered predictions, we conducted the following exploratory analyses. First, we examined how vocal parameters related to physical strength and body morphology. Second, we compared whether distinct parameters are used as cues for ratings on social dominance (i.e. being respected) and physical dominance (i.e. fighting ability), as they describe separate aspects of social evaluation⁹³. Third, we explored whether jitter and shimmer influence attractiveness and dominance perceptions, as these acoustic parameters seem to provide information on male body shape. Jitter and shimmer quantify cycle-to-cycle variation in f_0 and amplitude, respectively, and influence perceptions of voice roughness. Fourth, we conducted three mediation analyses: (1) a moderated mediation model to test whether f_0 mediates the relationship between vocalizers' testosterone levels (condition) and dominance ratings (perception), and whether this mediation is further moderated by cortisol, (2) a mediation model to test whether f_0 and P_f mediate the relationship between vocalizers' height and dominance ratings, and (3) a mediation model to test whether f_0 and P_f mediate the relationship between vocalizers' composite measure of size (extracted via factor analysis with varimax rotation) and dominance ratings. We conducted a separate mediation model for height, in addition to its inclusion in the factor analysis, as height has been shown to reflect good nutrition and low stress during development, as well as genetic predictors of immune function⁹⁴. Additionally, a recent study³¹ reported that f_0 mediated the relationship between height and physical dominance ratings in two separate samples. Finally, we conducted three meta-analyses to test: (1) the mediating effect of f_0 between height and dominance ratings, (2) whether cortisol and testosterone negatively interact to predict male f_0 , and (3) whether f_0 negatively predicts testosterone levels, especially among men with lower cortisol levels.

Design and methods

Participants. One hundred sixty-five heterosexual males participated in a study on testosterone reactivity and personality state changes, which was conducted at the University of Goettingen, Germany (for details, see⁹⁵). Each participant provided a standardized video recording, saliva samples, body morphology measurements, and handgrip as well as upper-body strength. Data from five individuals could not be used due to technical issues during video recording or because consent for further use of the video material was not given, resulting in a final sample of 160 males (mean age = 24.28, $SD = 3.25$ years). All participants were at least 18 years old. In a sensitivity power analysis using G*Power⁹⁶ this sample had sufficient power (>0.80) to detect an effect size of $r = +/ - 0.20$, assuming one-tailed alpha = 0.05. All procedures were in accordance with relevant guidelines and regulations, and received ethics approval from the local Ethics Committees at the University of Goettingen and the Pennsylvania State University. Informed consent was obtained from all subjects.

Voice recordings. Standardized video recordings were obtained using a Full-HD camera and Line6 Modell XD-V75 microphones. The participants were instructed to describe what is great about themselves, choosing three domains such as "friendship" or "success in studies/job" from a list of overall eight domains (for details, see⁹⁵). The video clips were cut to a length of 5 s, beginning 5 s after participants had begun to speak, and voice clips were extracted. Five seconds were chosen because vocal parameters usually show strong correlations across different recordings, independent of length and content^{88,97}, and both attractiveness and dominance ratings are stable and highly correlated across different recordings^{6,97}. Further, the use of relatively brief voice clips allowed us to avoid rater fatigue. The voice clips were analyzed using PRAAT software⁹⁸ (Version 6.0.36). The measures obtained were mean f_0 , the first four formant frequencies (F_1 – F_4), four measures of jitter and five measures of shimmer. Because both jitter (all $r_s > 0.83$, $ps < 0.001$) and shimmer measures (all $r_s > 0.56$, $ps < 0.001$) were highly intercorrelated, a standardized mean was calculated for each perturbation measure¹⁰. Additionally, P_f was computed for the first four formants⁶. Formants were measured at each glottal pulse using automated detection in PRAAT. Formant measurement across standardized speech samples produces highly similar results to measurement of individual vowels and averaging across these measurements⁶.

It should be noted that different methods of measuring formant structure are used across studies. Formant dispersion (D_f), for example, describes the distance between the highest (e.g., F_4) and lowest formants (e.g., F_1) measured³⁹. While D_f is commonly used, it has also been criticized especially for not using information about the middle formants (e.g., F_2 and F_3). Further, although D_f is theoretically dependent on body height, other measures of formant structure have shown stronger relations with body height^{6,33}. One of these measures is formant position (P_f) which describes the average standardized formant value for the first n formants (e.g., F_1 – F_4) and thus utilizes information of all formants measured⁶. Given these advantages of P_f over D_f , P_f was chosen as the relevant measure for formant structure in this study. For further discussion, see⁶.

Saliva samples. Based on previous studies^{99,100}, we controlled for circadian variation in participants' hormonal reactivity by collecting saliva samples only between 2 and 6 pm. Approximately 12–15 min after each participant arrived at the lab, he rinsed his mouth with water and provided at least 2 ml of saliva via passive drool through a straw, just prior to the video recording. The collected samples were immediately transported

to an ultra-low temperature freezer ($-80\text{ }^{\circ}\text{C}$), where salivary testosterone is expected to be stable for at least 36 months¹⁰¹. At the end of the data collection period (see⁹⁵ for details), saliva samples were shipped on dry ice to the Technical University of Dresden and analyzed using chemiluminescence-immuno-assays with high sensitivity (IBL International, Hamburg, Germany). The intra- and inter-assay coefficients (CVs) for cortisol are below 8% and for testosterone below 11%. Basal cortisol and testosterone outliers were identified and winsorized to 3 SDs¹⁰². To correct for skewness, we log₁₀-transformed both variables.

Body morphology and strength measurements. As this procedure was also reported in¹⁰³, procedural and methodological descriptions overlap. Participants were scanned three times using a Vitus Smart XXL 3D body scanner, running AnthroScan software (both Human Solutions GmbH, Kaiserslautern, Germany). Participants wore standardized tight underwear and were instructed to stand upright with legs hip-width apart, arms extended and held slightly away from the body, making a fist with thumbs showing forward, the head positioned in accordance with the Frankfurt Horizontal, and to breathe normally during the scanning process. Using AnthroScan's automatic measures (according to ISO 20685), we extracted muscularity-relevant body dimensions from the body scan: body volume, bust-chest girth, buttock girth, chest-to-hip ratio (CHR), forearm girth, lower limb ("leg") length-to-height ratio (LHR), shoulder-to-hip ratio (SHR), thigh girth, upper arm girth, waist girth, waist-to-chest ratio (WCR), and waist-to-hip ratio (WHR). An aggregate indicator of upper body size was calculated by averaging z -standardized shoulder width, bust-chest girth, and upper arm girth¹⁰⁴. Weight (in kg) was measured as part of the first body scanning process with the integrated SECA 635 scale (SECA, Hamburg, Germany). Body height (in cm) was measured twice using a stadiometer while participants stood barefoot, and the two values were averaged ($ICC=0.996$). Body-mass index (BMI) was calculated from average weight and height measures (kg/cm^2). Upper body and handgrip strength were measured using a hand dynamometer (Saehan SH5001). Each measurement was taken three times, starting with handgrip strength, for which participants were asked to use their dominant hand (88.2% used their right). As in¹⁰⁵, upper body strength was measured by having participants hold the dynamometer in front of their chest with both hands and press both handles toward the middle as strongly as possible. A composite strength measure was formed by averaging the maximum values for each of the three measures of handgrip and upper body strength ($ICCs$: 0.81 and 0.64, respectively).

Attractiveness and dominance ratings. In exchange for course credit, 120 men (mean age = 19.82, $SD=2.71$ years) and 120 women (mean age = 19.90, $SD=3.80$ years) participated in a rating study on short- and long-term attractiveness as well as social and physical dominance at the Pennsylvania State University. All raters were at least 18 years old. Raters were equipped with Sennheiser HD 280 Professional Headphones and seated at private workstations. Raters provided demographic data on age, gender, sexual orientation, and relationship status. To control for the influence of semantic content, we also asked raters to indicate their German language comprehension ("How well do you understand German?") on a 7-point Likert scale from 0 ("Not at All") to 6 ("Fluent"). Below, we report results with all participants, but excluding raters score 2 or higher ($n=26$) does not change results. Raters were then randomly assigned to one of four rating experiments, each asking for perceptions of either short-term attractiveness, long-term attractiveness, social dominance, or physical dominance of 160 randomly assigned voice files (for specific items see [Appendix A](#)). The voice file pool contained 320 voice samples that were taken from the 160 former targets before and after the competitive setting⁹⁵. Raters always rated both files of a target, but both recordings of the same individual were separated by at least ten other voice samples. However, only ratings of the recordings before the competition were used in the present study. To ensure that each file was rated 15 times by each sex, a file was removed from the pool of remaining files to be rated once this criterion was met. The only exception was long-term attractiveness, where one male rater dropped out because of technical issues. Because correlations between male and female ratings were high (all $r_s > 0.70$, $p_s < 0.001$), and intraclass correlations within each rating condition were at least satisfactory (all $ICCs > 0.76$, $p_s < 0.001$), mean scores were calculated.

Results

For tests of directed hypothesis one-tailed tests were used, and for exploratory tests two-tailed tests were used. Analyses were conducted using R¹⁰⁶.

Perceptions of Attractiveness and Dominance. *Attractiveness H1*) Predictions on negative linear and negative quadratic relationships between attractiveness ratings and mean f_o were supported. We found that f_o negatively linearly predicted both short-term and long-term attractiveness. Furthermore, we found significant negatively quadratic (inverted U-shaped) relationships between f_o and both short-term (Fig. 1a) and long-term attractiveness (Fig. 1b). Comparisons of linear and curvilinear models showed that the relationship between f_o and short-term attractiveness was significantly better described by the curvilinear model ($F_{2,157}=4.38$, $p=0.038$), while there was no significant difference between models for long-term attractiveness ($F_{2,157}=3.76$, $p=0.054$).

H2) Predictions of negative linear and negative quadratic relationships between attractiveness ratings and P_f were only partially supported. We found no significant linear relationships between P_f and either short-term or long-term attractiveness. While the non-linear relationship of P_f and short-term attractiveness was not significant (Fig. 2a), a significant negative quadratic relationship between P_f and long-term attractiveness emerged (Fig. 2b).

H3) The prediction of a stronger relationship between mean f_o and short-term, compared to long-term attractiveness ratings was supported. Although both attractiveness ratings were highly correlated ($r=0.82$, $p<0.001$), the relationship between f_o and short-term attractiveness was significantly stronger ($z=-2.06$, $p=0.020$) when comparing dependent correlation coefficients¹⁰⁸.

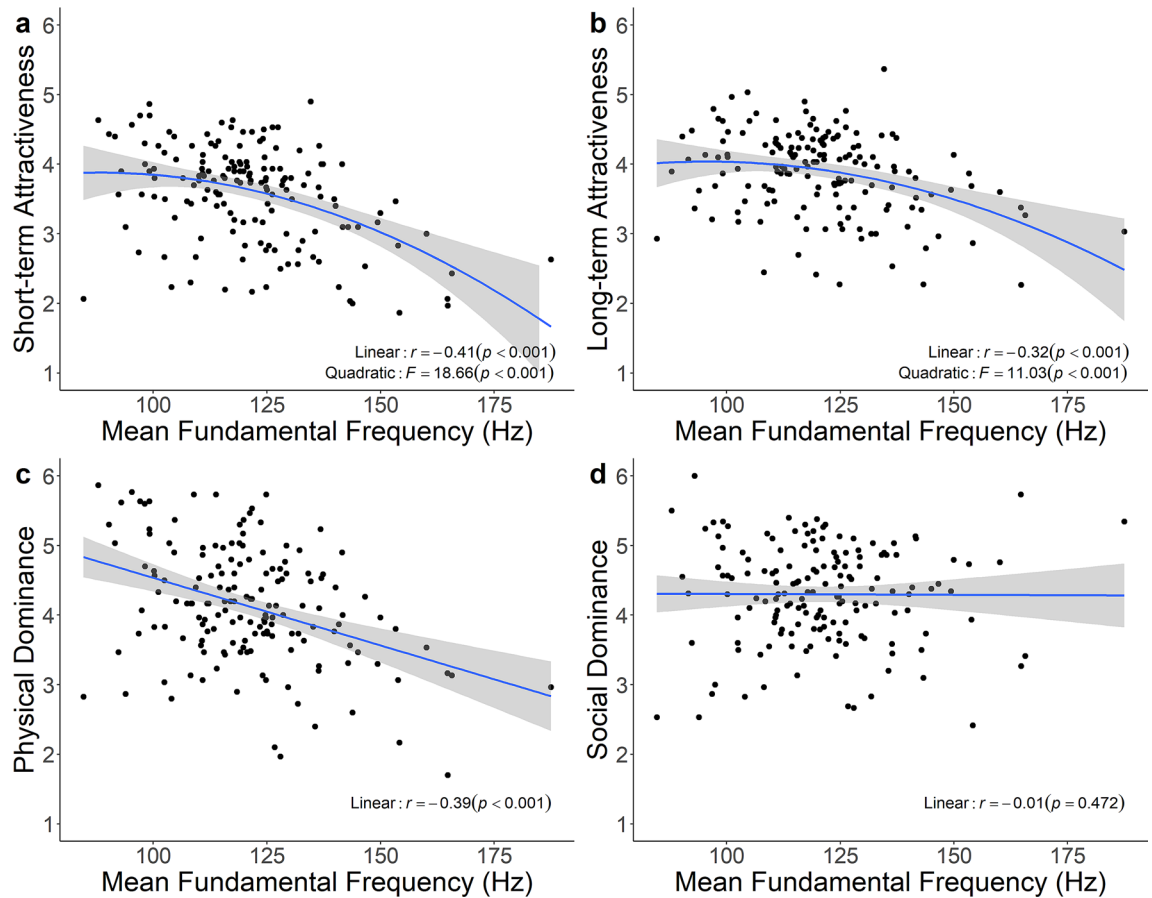


Figure 1. Relationships between male fundamental frequency (f_0) and perceptions. We observed negative curvilinear relationships between f_0 and (a) short-term attractiveness and (b) long-term attractiveness, (c) a negative linear relationship with physical dominance ratings, and (d) a non-significant relationship with social dominance ratings. All panels were plotted using the “ggplot2” package¹⁰⁷.

H4) The prediction of a stronger relationship between P_f and short-term, compared to long-term attractiveness ratings was supported; the relationship between P_f and short-term attractiveness was significantly stronger ($z = -2.00$, $p = 0.023$) when comparing dependent correlation coefficients.

Dominance H5) The prediction of a negative relationship between dominance ratings and mean f_0 was partially supported: f_0 negatively predicted physical dominance (Fig. 1c), but not social dominance ratings (Fig. 1d). **H6)** The prediction of a negative relationship between dominance ratings and P_f was supported. P_f negatively predicted perceptions of both physical (Fig. 2c) and social (Fig. 2d) dominance ratings.

Independent Predictors H7) The prediction that mean f_0 and P_f are independent predictors of attractiveness ratings was partially supported. When f_0 and P_f were included in a multiple regression ($F_{2,157} = 16.78$, $p < 0.001$, $R^2 = 0.17$), f_0 negatively predicted short-term attractiveness ($\beta = -0.40$, $p < 0.001$), but P_f did not ($\beta = -0.08$, $p = 0.132$). Similarly, f_0 negatively predicted long-term attractiveness ($\beta = -0.32$, $p < 0.001$) in a multiple regression ($F_{2,157} = 8.94$, $p < 0.001$, $R^2 = 0.09$), but P_f did not ($\beta = 0.01$, $p = 0.471$). Because the curvilinear relationship between long-term attractiveness and P_f was significant, we investigated whether the linear term of f_0 and the quadratic term of P_f were independent predictors of long-term attractiveness. Indeed, adding the quadratic term of P_f explained significantly more variance in long-term attractiveness ratings ($F_{2,157} = 3.15$, $p = 0.045$), with both predictors remaining significant. **H8)** The prediction that mean f_0 and P_f are independent predictors of dominance ratings was partially supported. Multiple regressions with f_0 and P_f as predictors ($F_{2,157} = 31.73$, $p < 0.001$, $R^2 = 0.28$) showed that both independently predicted physical dominance ($\beta = -0.35$, $p < 0.001$ for f_0 ; $\beta = -0.37$, $p < 0.001$ for P_f). For social dominance ($F_{2,157} = 5.12$, $p = 0.007$, $R^2 = 0.05$), P_f was a significant predictor ($\beta = -0.25$, $p < 0.001$), but f_0 was not ($\beta = 0.02$, $p = 0.391$).

Indirect measures of mate quality and formidability. *Testosterone, cortisol and f_0* Testosterone levels were not significantly related to f_0 ($r = -0.07$, $p = 0.18$). However, cortisol and testosterone interacted in predicting f_0 ($\beta = 0.16$, $p = 0.024$) (Fig. 3a). While these results do not support **H9)** a negative relationship between mean f_0 and testosterone, they supported **H10)** a negative relationship between mean f_0 and testosterone, which is attenuated by high baseline cortisol.

Body Morphology and P_f A significant relationship between P_f and body height was found ($r = -0.13$, $p = 0.046$), supporting **H11)**.

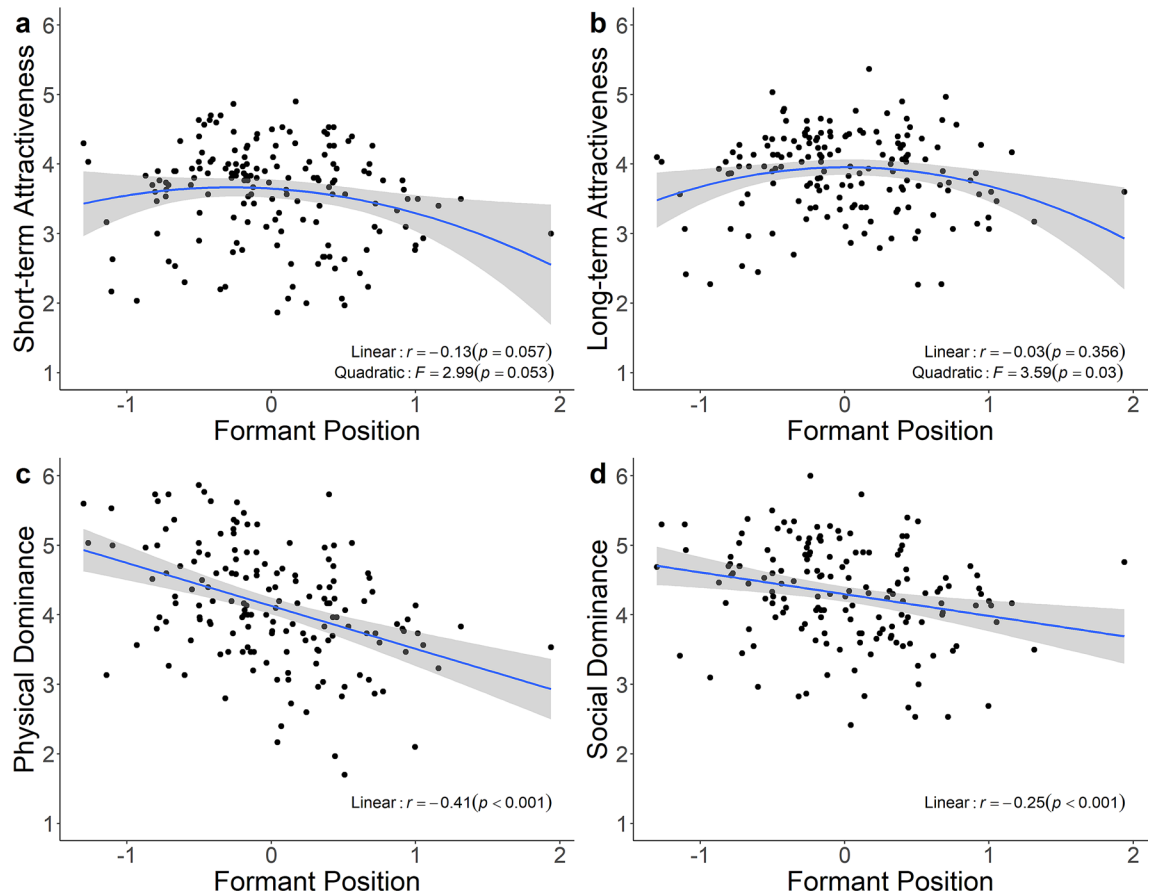


Figure 2. Relationships between male formant position (P_f) and perceptions. We observed negative curvilinear relationships between P_f and (a) short-term attractiveness and (b) long-term attractiveness, (c) a negative linear relationship with physical dominance ratings, and (d) social dominance ratings. All panels were plotted using the “ggplot2” package¹⁰⁷.

Exploratory analyses. Strength and P_f Additional exploratory analyses showed significant negative relationships between P_f and strength ($r = -0.25$, $p = 0.002$). Further, P_f was significantly correlated with multiple body morphology measures related to volume and mass (Table 2).

Perturbation measures, vocal perception and target parameters Pearson correlations showed significant negative relationships between shimmer and both social ($r = -0.31$, $p < 0.001$) and physical dominance ($r = -0.31$, $p < 0.001$). No significant relationships were found between shimmer and short-term ($r = -0.14$, $p = 0.076$) or long-term attractiveness ($r = -0.12$, $p = 0.122$). Jitter showed no significant relationship to any of the four ratings (all r s $< +/ - 0.11$, p s > 0.16). Moreover, the only significant relationship between perturbation measures and any of the target parameters was a significant negative correlation between shimmer and baseline cortisol ($r = -0.21$, $p = 0.006$). Multiple regressions with f_o , P_f , jitter and shimmer as predictors and all ratings as outcomes can be found in Tables S1–S4.

Mediation models In this analysis (model 7)¹¹¹, cortisol level was recoded into two categories (median split), and their interaction term was computed by multiplying testosterone levels with dichotomized cortisol category. In this model, we found that testosterone levels ($\beta = -0.09$; $p = 0.321$), cortisol category ($\beta = 0.07$; $p = 0.367$) and their interaction term ($\beta = 0.135$; $p = 0.119$) did not predict f_o . Adjusting for P_f ($\beta = -0.39$; $p < 0.001$), testosterone ($\beta = 0.15$; $p = 0.023$) and f_o ($\beta = -0.34$; $p < 0.001$) significantly predicted physical dominance ratings. The indirect effect of testosterone on dominance ratings via f_o was not significant ($\beta = 0.06$; $p = 0.344$), and no significant indirect effect was observed among men with lower cortisol ($\beta = 0.04$; $p = 0.227$), or men with higher cortisol levels ($\beta = 0.02$; $p = 0.832$).

We ran two additional mediation models: (1) f_o and P_f were entered as mediators between height and physical dominance ratings, (2) f_o and P_f were entered as mediators between physical strength and dominance ratings. A composite measure of physical size was extracted from a factor analysis (Fig. 4d) on the following body morphology measures that significantly correlated with P_f (Table 2): height, weight, body volume, bust-chest girth, buttock girth, forearm girth, physical strength, thigh girth, upper body size, upper arm girth, and waist girth. In model 1, f_o and P_f were entered as mediators between height and physical dominance ratings (Fig. 4a). Neither f_o nor P_f was a significant mediator. In model 2, we found evidence that P_f mediated the relationship between physical strength condition and physical dominance ratings (Fig. 4b).

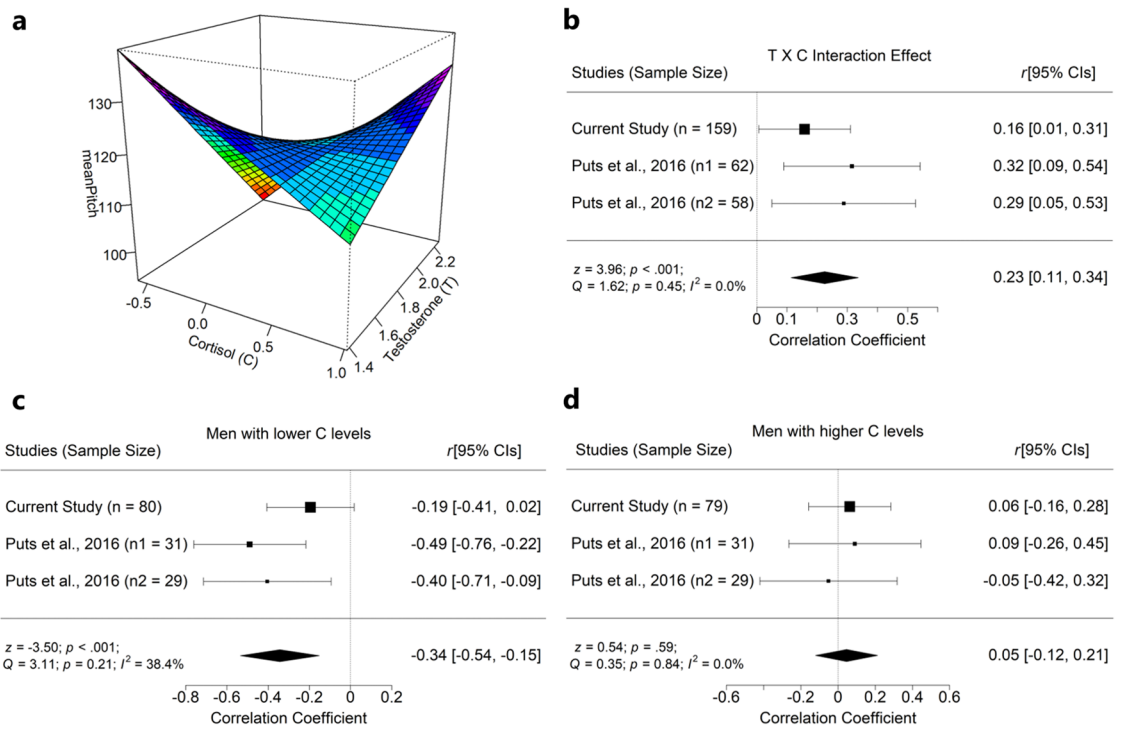


Figure 3. Negative interaction between testosterone and cortisol on male fundamental frequency (f_0). **(a)** A combination of higher testosterone and lower cortisol levels predict lower male f_0 in this study. **(b)** A meta-analysis on the interaction effects across studies, using a random-effects model yielded a significant overall effect. Follow-up meta-analyses on simple slopes of **(c)** lower cortisol levels yielded a significant negative relationship between testosterone and f_0 , and **(d)** higher cortisol levels yielded null results. Panel b was plotted via the “rsm” package¹⁰⁹, and meta-analyses were conducted via the “metaphor” package¹¹⁰.

| Variable | M | SD | r | 95% CI |
|----------------------------------|--------|-------|---------|--------------|
| BMI | 23.98 | 3.83 | -.23*** | [-.37, -.08] |
| Body volume | 79.88 | 14.03 | -.27*** | [-.41, -.12] |
| Bust-chest girth | 101.67 | 8.81 | -.29*** | [-.43, -.14] |
| Buttock girth | 100.18 | 7.25 | -.26*** | [-.40, -.11] |
| Forearm girth | 27.00 | 1.93 | -.28*** | [-.42, -.13] |
| Physical strength | 48.40 | 7.99 | -.25** | [-.39, -.09] |
| Thigh girth | 57.58 | 4.97 | -.22** | [-.37, -.07] |
| Upper body size | 56.96 | 4.13 | -.31*** | [-.44, -.16] |
| Upper arm girth | 30.20 | 2.67 | -.25** | [-.39, -.09] |
| Waist girth | 84.63 | 9.86 | -.24** | [-.39, -.09] |
| Weight | 78.68 | 13.96 | -.27*** | [-.41, -.12] |
| Chest-to-hip ratio (CHR) | 1.02 | 0.05 | -.13 | [-.28, .02] |
| Waist-to-chest ratio (WCR) | 1.21 | 0.07 | .03 | [-.13, .18] |
| Waist-to-hip ratio (WHR) | 0.84 | 0.05 | -.15 | [-.30, .00] |
| Leg length-to-height ratio (LHR) | 0.40 | 0.01 | .12 | [-.03, .27] |
| Shoulder-to-hip ratio (SHR) | 0.39 | 0.02 | .08 | [-.08, .23] |

Table 2. Means, standard deviations, and correlations of body morphology measures with P_f . *M* and *SD* are used to represent mean and standard deviation. Values in square brackets indicate the confidence interval for each correlation. ** indicates $p < .01$; *** indicates $p < .001$.

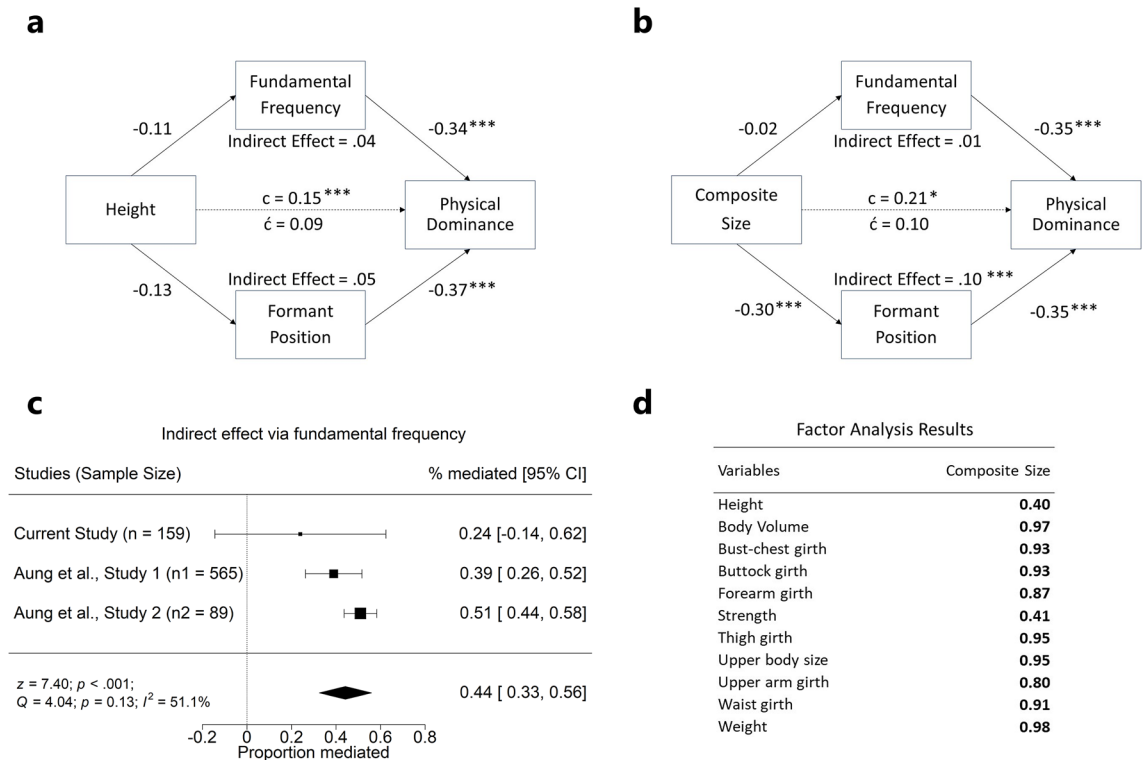


Figure 4. Male fundamental frequency (f_o) and formant position (P_f) as mediators of vocalizers’ condition and perceivers’ ratings. **(a)** Although height predicted physical dominance ratings, f_o and P_f did not mediate this relationship. **(b)** P_f but not f_o , significantly mediated the relationship between composite size and physical dominance ratings. **(c)** Although f_o was not found to be a significant mediator between height and physical dominance ratings in the present study, a meta-analysis using a random-effects model indicated a significant mediating effect, with f_o mediating 44% of the relationship between height and physical dominance. Proportion mediated lower than 0 indicates the suppression effect of a mediating variable. In addition, the current study used mean dominance ratings as the primary unit of analyses for calculating proportion mediated, whereas Aung et al., Study 1 (n = 8,103 observations) and Study 2 (n = 6,586 observations) used individual ratings. **(d)** Using the “nFactors” package¹¹² and rotated factors with Varimax method using the “psych” package¹¹³, we reduced the set of size related measures into one dimensional factor (n = 1), which we labelled “composite size”, via principal axis factoring analysis. ***p < .001.

Meta-analyses We combined results of the present study with prior results³¹ in a meta-analysis to assess the strength of the mediating effect of f_o on the relationship between height and perceptions of physical dominance. We found a significant overall mediating effect of f_o , independent of P_f (Fig. 4c); f_o mediated about 44% the relationship between height and physical dominance ratings.

We also conducted a meta-analysis of the interaction of testosterone and cortisol in predicting f_o . For this analysis, the *t*-value and degrees of freedom (*df*) of the overall interaction effect were transformed into a correlation¹¹⁴. The effect of the testosterone and cortisol interaction on male f_o (*k* = 3, *n* = 279) was significant: *r* = 0.23, *p* = 0.001, 95% CI [0.12, 0.34] (Fig. 3b). In follow-up analyses, the relationship between testosterone and f_o was significant in men with low cortisol levels (Fig. 3c), but not in those with high cortisol levels (Fig. 3d).

Finally, Fig. 5 provides a lens model¹¹⁵ overview of the key relations between perceptions, vocal cues and target parameters found in this study.

Discussion

We investigated the role of vocal parameters in perceptions of male attractiveness and found that f_o was the strongest predictor of short- and long-term attractiveness among the vocal parameters measured (P_f , shimmer, and jitter). Consistent with previous studies^{11,64}, the relationship between f_o and male vocal attractiveness was both negatively linear and negatively curvilinear, the latter suggesting that women’s voice preferences may reflect a tradeoff between the potential genetic or other benefits versus the potential costs of mating with masculine males⁶. Such costs may include lower investment and perhaps risk of interpersonal violence. Low male f_o has previously been linked to sexual infidelity^{42,88,89,116}, and several lines of evidence suggest that phenotypic masculinity—and vocal masculinity in particular—indicate threat potential not only to same-sex competitors but also to potential mates. For example, images of male-on-female violence disrupted U.S. women’s preferences for both masculine voices and faces¹¹⁷, and Colombian women with perceptions of greater local domestic violence

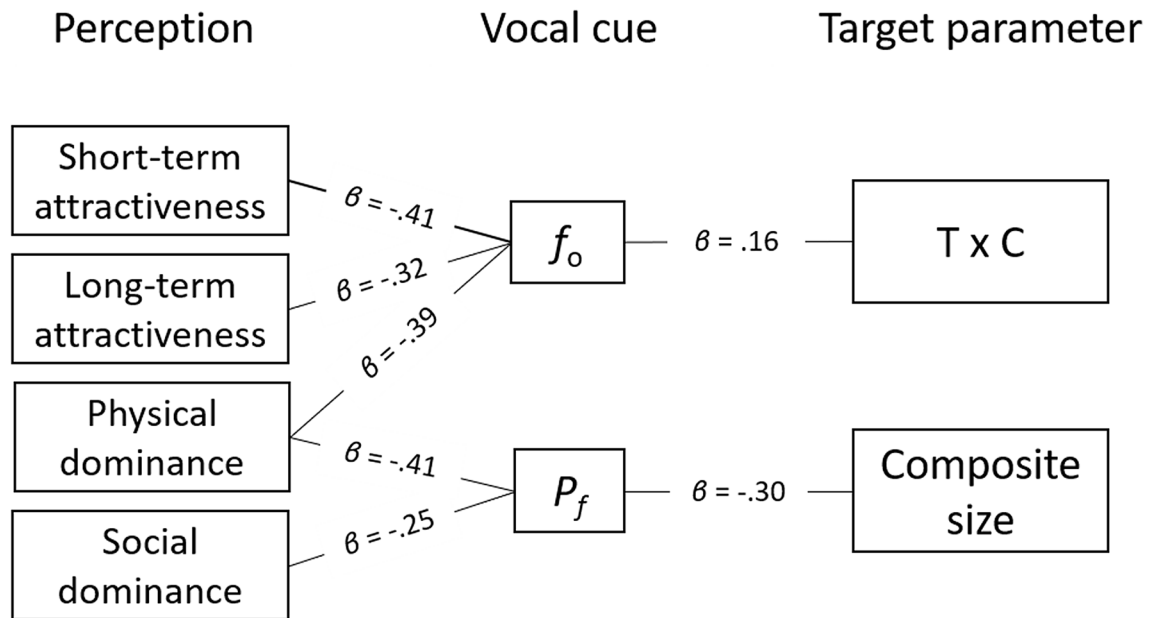


Figure 5. Lens model overview of the study results. Connections indicate significant relations ($p < .05$).

preferred less masculine male faces¹¹⁸. In another study, Filipino women who were younger and rated themselves as less attractive tended to prefer feminized male f_o , again suggesting that women's f_o preferences may in part reflect their own perceived vulnerability⁶². In our data, f_o was a stronger predictor of short-term than long-term attractiveness, once again supporting the notion of a mate choice trade-off in which putative indicators of genetic fitness are prioritized in short-term contexts, and expected investment and fidelity are prioritized in long-term contexts⁹⁰.

Although P_f predicted strength and body morphology in our study and predicted ratings of attractiveness in some prior studies^{13,21}, it did not predict attractiveness in another large sample⁵ and was unrelated to short-term attractiveness and only weakly negatively curvilinearly linked to long-term attractiveness in the present study. These lines of evidence suggest that the information provided by formant frequencies may be less relevant to mate quality than that provided by f_o . By contrast, shimmer negatively predicted both short- and long-term attractiveness ratings. Shimmer is utilized to assess vocal quality in clinical contexts, such that pathological voices show higher shimmer levels than those of healthy individuals^{119–121}; however, a composite of shimmer and harmonics-to-noise ratio (which were highly correlated) showed no relationship to dominance or attractiveness perception in a recent study⁵. These divergent findings may be explained by the fact that the latter study used voice samples in which male individuals read a standardized voice passage, while our study used more natural but less standardized stimuli that might have been influenced more strongly by the speaker's affective state.

Importantly, a Fisherian mate choice model via runaway sexual selection has also been suggested as a possible driver favoring low male f_o ^{14,122}. A Fisherian model would suggest that female choice primarily drives and exaggerates the evolution of male traits; hence, the model predicts that females prefer males with the lowest f_o . However, evidence from the current study and previous studies^{15,62,65} suggests a general preference for lower f_o by women, but also a relatively stronger negative linear relationship between f_o and dominance perceptions by men across studies²⁹.

While f_o predicted both short- and long-term attractiveness, it predicted physical dominance but not social dominance, in line with previous studies^{13,123}. P_f and shimmer were linked to both social and physical dominance ratings. A possible explanation for this pattern of results is that social dominance is influenced less by threat potential and more by other qualities, such as competence, communication and cooperation skills, or leadership qualities. These attributes might be more strongly associated with P_f and shimmer than with f_o .

The other aim of this study was to explore whether attention to vocal cues is adaptive by investigating the information content of acoustic parameters. We replicated a negative relationship between P_f and height³³ and found that P_f negatively predicted strength and several body morphology measures. Men with lower P_f were taller, stronger, and had larger bodies in general. Further, our mediation analysis indicated that P_f , independently of f_o , mediated the relationship between a composite measure of body size and physical dominance ratings.

Importantly, baseline cortisol and testosterone levels interacted in predicting f_o , such that testosterone levels more strongly negatively predicted f_o as cortisol levels decreased across participants. When we entered the interaction term between testosterone and median-split cortisol levels into our exploratory moderated mediation analyses, the interaction effect became non-significant, likely due to reduced statistical power¹²⁴ from dichotomizing a continuous variable (cortisol). Nevertheless, the overall interaction between testosterone and cortisol in predicting male f_o was confirmed in a meta-analysis (Fig. 3b). Male f_o was negatively correlated with

testosterone when cortisol was low, whereas no significant relationship was observed between male f_0 and testosterone when cortisol was high (Fig. 3c). These patterns of relationships may help clarify why dose-dependent effects of androgen levels on the intensity of elaborate male traits are sometimes undetected¹²⁵, and why f_0 is only weakly correlated with testosterone when cortisol is not considered. Across a variety of species, testosterone and cortisol are linked to measures of physical condition, including disease, stress, and diet¹²⁶. The interaction between testosterone and cortisol, in particular, has been tied to immune function in birds¹²⁷, but the functional and behavioral correlates of this hormonal interaction in humans are not yet clear^{36,128}, and most studies are arguably underpowered. Further, a recent meta-analysis found only modest support for an interactive relationship between testosterone and cortisol in predicting status-relevant behavior (e.g., dominance & risk taking) and suggested that this association could be driven by publication bias and flexibility in data analysis^{129,130}. Although only one paper⁵ besides the current one has reported the specific interaction effect of testosterone and cortisol on male f_0 , the meta-analysis reported here suggests that the interaction is robust.

There is widespread agreement^{5,11,27,40,46} that low male f_0 evolved to exaggerate apparent size by leveraging a predisposition to perceive low frequencies as emanating from large sound sources. Phylogenetic reconstruction suggests that relatively male f_0 evolved in the common ancestor of the catarrhine primates after their divergence from platyrrhines approximately 43.5mya⁵. Given the weak correspondence between f_0 and body size, some have argued that f_0 is purely deceptive and is not an honest indicator of physical dominance^{27,28,131}. Others have suggested that f_0 may reliably correlate with other salient speaker characteristics such as status, threat, and dominance, and that these dimensions may overlap with, and hence intrude onto impressions of, size⁴⁶. Our results better comport with the latter possibility. Indeed, relatively low male f_0 tends to be lost in primate species in which male-male mating competition is reduced, suggesting that there are costs associated with low f_0 that cause this trait to be selected against when compensatory benefits are absent.

Deference to males with low f_0 is demonstrably costly in humans in terms of social status, mates, and reproduction, and thus attention to f_0 would seemingly be selected against if f_0 did not provide valid information about male condition³⁰. However, this does not mean that f_0 is cheat-proof, or that the assessment of condition or formidability from f_0 is largely accurate. Honest signals are often corrupted into conventional signals where cheating is common because the assessment of the signal itself is costly to the receiver¹³². Although we did not find support for the cortisol-moderated mediation role of f_0 between testosterone levels and physical dominance ratings in the present sample, this may be explained by reduced statistical power due to dichotomized cortisol levels and reduced sample sizes for testing two separate indirect effects. Indeed, we found a strong meta-analytic support for an overall interaction between testosterone and cortisol in predicting male f_0 , suggesting that f_0 conveys underlying endocrine state, if imprecisely, and lower male f_0 has consistently been shown to predict perceptions of physical dominance across multiple studies. Likewise, a recent study³¹ reported that f_0 mediated the relationship between developmental condition (measured via height) and physical dominance ratings in two separate samples with different types of vocal stimuli. Although we did not find that f_0 significantly mediated the relationship between height and physical dominance ratings in our data, our meta-analysis suggests that f_0 mediates about 44% of the relationship between height and physical dominance ratings. Collectively, our findings support the hypothesis that, while the correlation between f_0 and underlying quality is imperfect, f_0 might be utilized as one of many cues for assessing competitors and potential mates²⁹ because it communicates the quality of the signaler significantly better than chance^{132,133}.

Shimmer also negatively predicted social and physical dominance ratings, as well as lower cortisol levels. The latter finding is consistent with prior evidence that shimmer is reduced when stress is induced experimentally or when the speaker is under high tension¹³⁴. However, the other perturbation measure, jitter, showed no such associations. Future research should continue to explore the relevance of jitter and shimmer to human sexual selection (see also), as they have been shown to be associated with pathological voice quality¹²⁰ and body shape in men⁴¹ and might therefore be relevant in contexts of sexual selection.

One limitation with our study is that we tested only hypotheses associated with receiver-independent costs and did not consider receiver-dependent costs associated with attention to male f_0 . Some^{135,136} have suggested that additional mechanisms that incorporate receiver-dependent costs are required to ensure signal honesty. For example, under a mating-motive priming condition, male voices with low f_0 enhanced recognition for men with high threat potential¹³⁵ and elicited aggressive cognitions and intent in men who perceived themselves to be more dominant and stronger¹³⁶. Future studies should investigate the extent to which receiver-dependent and independent costs are needed in ensuring the signal honesty of low f_0 in cross-cultural contexts.

Following suggestions by Lakens¹³⁷, we used one-sided significance tests for preregistered directional hypotheses. The only result influenced by this decision is the relation between P_f and height, which would be non-significant using a two-sided test. However, we note that meta-analytic findings³³ suggest a robust link between P_f and height, and the lack of a significant relation in this particular study is likely due to a lack of statistical power. Thus, also our conclusions remain highly similar when two-sided tests are used.

Conclusion

Vocal parameters were linked to hormone levels, as well as body morphology and physical strength, and appear to be used for judgements relevant to intrasexual competition and intersexual mate choice. The present study thus provides evidence that natural interindividual variation in men's vocal parameters influences judgements of attractiveness and dominance because these parameters provide valid information about speakers' underlying condition.

Appendix A

Demographic Questions for Raters

What is your sex? (Female/Male)

How old are you? (text box)

Relationship status: (Single, Dating, Cohabiting/Married)

Which best describes your sexual attractions?

“I am attracted...”

(only to females mostly to females equally to males and females mostly to males only to males)

How well do you understand German? (7-point Likert Scale: 0 Not at All - 6 Fluent)

Voice Rating Questions

1. **Short-term attractiveness:** How attractive does the speaker sound for a short-term, uncommitted romantic relationship?
 - a. 7-point Likert Scale (-3 very unattractive - 3+ very attractive)
2. **Long-term attractiveness:** How attractive does the speaker sound for a long-term, committed romantic relationship?
 - a. 7-point Likert Scale (-3 very unattractive - 3+ very attractive)
3. **Social dominance:** How respected does the speaker sound?
 - a. 7-point Likert Scale (-3 very unrespected - 3+ very respected)
4. **Physical dominance:** How likely would the speaker win a physical fight against an average college student?
 - a. 7-point Likert Scale (-3 very unlikely - 3+ very likely)

Received: 21 February 2020; Accepted: 18 November 2020

Published online: 04 December 2020

References

1. Cade, W. H. & Cade, E. S. Male mating success, calling and searching behaviour at high and low densities in the field cricket *Gryllus integer*. *Anim. Behav.* **43**, 49–56 (1992).
2. Fischer, J., Kitchen, D. M., Seyfarth, R. M. & Cheney, D. L. Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behav. Ecol. Sociobiol.* **56**, 140–148 (2004).
3. Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E. & Theunissen, F. E. What the hyena’s laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecol.* **10**, 9 (2010).
4. Mitani, J. C. Sexual selection and adult male orangutan long calls. *Anim. Behav.* **33**, 272–283 (1985).
5. Puts, D. A. *et al.* Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proc. R. Soc. B* **283**, 20152830 (2016).
6. Puts, D. A., Apicella, C. L. & Cárdenas, R. A. Masculine voices signal men’s threat potential in forager and industrial societies. *Proc. R. Soc. B Biol. Sci.* **279**, 601–609 (2012).
7. Ptacek, P. H. & Sander, E. K. Age recognition from voice. *J. Speech Hear. Res.* **9**, 273–277 (1966).
8. Rezlescu, C. *et al.* Dominant voices and attractive faces: the contribution of visual and auditory information to integrated person impressions. *J. Nonverbal Behav.* **39**, 355–370 (2015).
9. Sell, A. *et al.* Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. B Biol. Sci.* **277**, 3509–3518 (2010).
10. Jünger, J. *et al.* Do women’s preferences for masculine voices shift across the ovulatory cycle?. *Horm. Behav.* **106**, 122–134 (2018).
11. Puts, D. A., Jones, B. C. & DeBruine, L. M. Sexual selection on human faces and voices. *J. Sex Res.* **49**, 227–243 (2012).
12. Cartei, V., Bond, R. & Reby, D. What makes a voice masculine: physiological and acoustical correlates of women’s ratings of men’s vocal masculinity. *Horm. Behav.* **66**, 569–576 (2014).
13. Hodges-Simeon, C. R., Gaulin, S. J. C. & Puts, D. A. Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* **21**, 406–427 (2010).

14. Puts, D. A. Mating context and menstrual phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* **26**, 388–397 (2005).
15. Rosenfield, K. A., Sorokowska, A., Sorokowski, P. & Puts, D. A. Sexual selection for low male voice pitch among Amazonian forager-horticulturalists. *Evol. Hum. Behav.* **41**, 3–11 (2019).
16. Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J. & Penke, L. The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum. Behav.* **39**, 424–436 (2018).
17. Apicella, C. L., Feinberg, D. R. & Marlowe, F. W. Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett.* **3**, 682–684 (2007).
18. Smith, K. M., Olkhov, Y. M., Puts, D. A. & Apicella, C. L. Hadza men with lower voice pitch have a better hunting reputation. *Evol. Psychol.* **15**, 1474704917740466 (2017).
19. Atkinson, J. *et al.* Voice and handgrip strength predict reproductive success in a group of indigenous African females. *PLoS ONE* **7**, e41811 (2012).
20. Puts, D. A., Hodges, C. R., Cárdenas, R. A. & Gaulin, S. J. C. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* **28**, 340–344 (2007).
21. Hodges-Simeon, C. R., Gaulin, S. J. C. & Puts, D. A. Voice correlates of mating success in men: examining “contests” versus “mate choice” modes of sexual selection. *Arch. Sex. Behav.* **40**, 551–557 (2011).
22. Zahavi, A. Mate selection—a selection for a handicap. *J. Theor. Biol.* **53**, 205–214 (1975).
23. Grafen, A. Biological signals as handicaps. *J. Theor. Biol.* **144**, 517–546 (1990).
24. Higham, J. P. How does honest costly signaling work?. *Behav. Ecol.* **25**, 8–11 (2014).
25. Penn, D. J. & Számadó, S. The handicap principle: how an erroneous hypothesis became a scientific principle. *Biol. Rev.* **95**, 267–290 (2020).
26. Reid, S. A., Zhang, J., Anderson, G. L. & Keblusek, L. Costly signaling in human communication. in *The Handbook of Communication Science and Biology* (eds. Floyd, K. & Weber, R.) 50–62 (Routledge, 2020).
27. Armstrong, M. M., Lee, A. J. & Feinberg, D. R. A house of cards: bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Anim. Behav.* **147**, 43–51 (2019).
28. Feinberg, D. R., Jones, B. C. & Armstrong, M. M. No evidence that men's voice pitch signals formidability. *Trends Ecol. Evol.* **34**, 190–192 (2019).
29. Aung, T. & Puts, D. Voice pitch: a window into the communication of social power. *Curr. Opin. Psychol.* **33**, 154–161 (2019).
30. Puts, D. A. & Aung, T. Does men's voice pitch signal formidability? A reply to feinberg. *Trends Ecol. Evol.* **34**, 189–190 (2019).
31. Aung, T., Rosenfield, K. A. & Puts, D. Male voice pitch mediates the relationship between objective and perceived formidability. *Evol. Hum. Behav.* (2020).
32. Han, C. *et al.* No clear evidence for correlations between handgrip strength and sexually dimorphic acoustic properties of voices. *Am. J. Hum. Biol.* **0**, e23178.
33. Pisanski, K. *et al.* Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* **95**, 89–99 (2014).
34. Dabbs, J. M. & Mallinger, A. High testosterone levels predict low voice pitch among men. *Personal. Individ. Differ.* **27**, 801–804 (1999).
35. Evans, S., Neave, N., Wakelin, D. & Hamilton, C. The relationship between testosterone and vocal frequencies in human males. *Physiol. Behav.* **93**, 783–788 (2008).
36. Rantala, M. J. *et al.* Evidence for the stress-linked immunocompetence handicap hypothesis in humans. *Nat. Commun.* **3**, 694 (2012).
37. Arnocky, S., Hodges-Simeon, C. R., Ouellette, D. & Albert, G. Do men with more masculine voices have better immunocompetence?. *Evol. Hum. Behav.* **39**, 602–610 (2018).
38. Albert, G., Arnocky, S., Puts, D. A. & Hodges-Simeon, C. R. Can listeners assess men's self-reported health from their voice? *Evol. Hum. Behav.* (2020).
39. Fitch, W. T. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.* **102**, 1213–1222 (1997).
40. Fitch, W. T. The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267 (2000).
41. Pisanski, K. *et al.* Voice parameters predict sex-specific body morphology in men and women. *Anim. Behav.* **112**, 13–22 (2016).
42. Schild, C., Stern, J. & Zettler, I. Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behav. Ecol.* **31**, 164–175 (2019).
43. Collins, S. A. & Missing, C. Vocal and visual attractiveness are related in women. *Anim. Behav.* **65**, 997–1004 (2003).
44. Raine, J., Pisanski, K., Bond, R., Simner, J. & Reby, D. Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech. *PLoS ONE* **14**, e0213034 (2019).
45. Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J. & Reby, D. Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* **4**, 273–280 (2018).
46. Rendall, D., Vokey, J. R. & Nemeth, C. Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size. *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 1208–1219 (2007).
47. Šebesta, P. *et al.* Voices of Africa: acoustic predictors of human male vocal attractiveness. *Anim. Behav.* **127**, 205–211 (2017).
48. Šebesta, P., Třebický, V., Fialová, J. & Havlíček, J. Roar of a champion: loudness and voice pitch predict perceived fighting ability but not success in mma fighters. *Front. Psychol.* **10**, 859 (2019).
49. Simmons, L. W., Peters, M. & Rhodes, G. Low pitched voices are perceived as masculine and attractive but do they predict semen quality in men?. *PLoS ONE* **6**, e29271 (2011).
50. Valentová, J. V. *et al.* 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 (2019).
51. Feinberg, D. R., DeBruine, L. M., Jones, B. C. & Perrett, D. I. The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception* **37**, 615–623 (2008).
52. Babel, M., McGuire, G. & King, J. Towards a more nuanced view of vocal attractiveness. *PLoS ONE* **9**, e88616 (2014).
53. Gregory, S. W., Dagan, K. & Webster, S. Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *J. Nonverbal Behav.* **21**, 23–43 (1997).
54. Knowles, K. K. & Little, A. C. Vocal fundamental and formant frequencies affect perceptions of speaker cooperativeness. *Q. J. Exp. Psychol.* **69**, 1657–1675 (2016).
55. Michalsky, J. & Schoormann, H. Pitch convergence as an effect of perceived attractiveness and likability. in *Interspeech 2017* 2253–2256 (ISCA, 2017).
56. Pisanski, K. & Rendall, D. The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *J. Acoust. Soc. Am.* **129**, 2201–2212 (2011).
57. Pisanski, K., Mishra, S. & Rendall, D. The evolved psychology of voice: evaluating interrelationships in listeners' assessments of the size, masculinity, and attractiveness of unseen speakers. *Evol. Hum. Behav.* **33**, 509–519 (2012).
58. Sorokowski, P. *et al.* Voice of authority: professionals lower their vocal frequencies when giving expert advice. *J. Nonverbal Behav.* **43**, 257–269 (2019).
59. Valentová, J., Roberts, S. C. & Havlíček, J. Preferences for facial and vocal masculinity in homosexual men: the role of relationship status, sexual restrictiveness, and self-perceived masculinity. *Perception* **42**, 187–197 (2013).

60. Hill, A. K. *et al.* Are there vocal cues to human developmental stability? Relationships between facial fluctuating asymmetry and voice attractiveness. *Evol. Hum. Behav.* **38**, 249–258 (2017).
61. Wolff, S. E. & Puts, D. A. Vocal masculinity is a robust dominance signal in men. *Behav. Ecol. Sociobiol.* **64**, 1673–1683 (2010).
62. Shirazi, T. N., Puts, D. A. & Escasa-Dorne, M. J. Filipino women's preferences for male voice pitch: intra-individual, life history, and hormonal predictors. *Adapt. Hum. Behav. Physiol.* **4**, 188–206 (2018).
63. Re, D. E., O'Connor, J. J. M., Bennett, P. J. & Feinberg, D. R. Preferences for very low and very high voice pitch in humans. *PLoS ONE* **7**, e32719 (2012).
64. Saxton, T. K., Mackey, L. L., McCarty, K. & Neave, N. A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. *Behav. Ecol.* **27**, 512–519 (2016).
65. Apicella, C. L. & Feinberg, D. R. Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proc. R. Soc. B Biol. Sci.* **276**, 1077–1082 (2009).
66. Borkowska, B. & Pawlowski, B. Female voice frequency in the context of dominance and attractiveness perception. *Anim. Behav.* **82**, 55–59 (2011).
67. Bruckert, L. *et al.* Vocal attractiveness increases by averaging. *Curr. Biol.* **20**, 116–120 (2010).
68. Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M. & Perrett, D. I. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim. Behav.* **69**, 561–568 (2005).
69. Feinberg, D. R. *et al.* Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Horm. Behav.* **49**, 215–222 (2006).
70. Feinberg, D. R., DeBruine, L. M., Jones, B. C. & Little, A. C. Correlated preferences for men's facial and vocal masculinity. *Evol. Hum. Behav.* **29**, 233–241 (2008).
71. Feinberg, D. R. *et al.* Integrating fundamental and formant frequencies in women's preferences for men's voices. *Behav. Ecol.* **22**, 1320–1325 (2011).
72. Fraccaro, P. J. *et al.* Faking it: deliberately altered voice pitch and vocal attractiveness. *Anim. Behav.* **85**, 127–136 (2013).
73. Hughes, S. M., Mogilski, J. K. & Harrison, M. A. The perception and parameters of intentional voice manipulation. *J. Nonverbal Behav.* **38**, 107–127 (2014).
74. Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C. & Vukovic, J. A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Anim. Behav.* **79**, 57–62 (2010).
75. Klofstad Casey, A., Anderson Rindy, C. & Peters, S. 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 (2012).
76. Leaderbrand, K., Dekam, J., Morey, A. & Tuma, L. The effects of voice pitch on perceptions of attractiveness: do you sound hot or not. *Winona State Univ. Psychol. Stud. J.* (2008).
77. O'Connor, J. J. M. *et al.* Female preferences for male vocal and facial masculinity in videos. *Ethology* **118**, 321–330 (2012).
78. Puts, D. A., Gaulin, S. J. C. & Verdolini, K. Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* **27**, 283–296 (2006).
79. Puts, D. A., Barndt, J. L., Welling, L. L. M., Dawood, K. & Burriss, R. P. Intrasexual competition among women: vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personal. Individ. Differ.* **50**, 111–115 (2011).
80. Riding, D., Lonsdale, D. & Brown, B. The effects of average fundamental frequency and variance of fundamental frequency on male vocal attractiveness to women. *J. Nonverbal Behav.* **30**, 55–61 (2006).
81. Suire, A., Raymond, M. & Barkat-Defradas, M. Male vocal quality and its relation to females' preferences. *Evol. Psychol.* **17**, 1474704919874675 (2019).
82. Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C. & Feinberg, D. R. Voice pitch influences voting behavior. *Evol. Hum. Behav.* **33**, 210–216 (2012).
83. Vukovic, J. *et al.* Variation in perceptions of physical dominance and trustworthiness predicts individual differences in the effect of relationship context on women's preferences for masculine pitch in men's voices. *Br. J. Psychol.* **102**, 37–48 (2011).
84. Watkins, C. D. *et al.* Taller men are less sensitive to cues of dominance in other men. *Behav. Ecol.* **21**, 943–947 (2010).
85. Xu, Y., Lee, A., Wu, W.-L., Liu, X. & Birkholz, P. Human vocal attractiveness as signaled by body size projection. *PLoS ONE* **8**, e62397 (2013).
86. O'Connor, J. J. M., Re, D. E. & Feinberg, D. R. Voice pitch influences perceptions of sexual infidelity. *Evol. Psychol.* **9**, 147470491100900100 (2011).
87. O'Connor, J. J. M. & Barclay, P. The influence of voice pitch on perceptions of trustworthiness across social contexts. *Evol. Hum. Behav.* **38**, 506–512 (2017).
88. Schild, C., Stern, J., Penke, L. & Zettler, I. Voice pitch—a valid indicator of one's unfaithfulness in committed relationships? *Adapt. Hum. Behav. Physiol.* <https://doi.org/10.1007/s40750-020-00154-0> (2020).
89. O'Connor, J. J. M., Pisanski, K., Tigue, C. C., Fraccaro, P. J. & Feinberg, D. R. Perceptions of infidelity risk predict women's preferences for low male voice pitch in short-term over long-term relationship contexts. *Personal. Individ. Differ.* **56**, 73–77 (2014).
90. Gangestad, S. W. & Simpson, J. A. The evolution of human mating: trade-offs and strategic pluralism. *Behav. Brain Sci.* **23**, 573–587 (2000).
91. Titze, I. R. *Principles of Voice Production* (Prentice Hall, Upper Saddle River, 1994).
92. Stulp, G., Pollet, T. V., Verhulst, S. & Buunk, A. P. A curvilinear effect of height on reproductive success in human males. *Behav. Ecol. Sociobiol.* **66**, 375–384 (2012).
93. Maner, J. K. & Case, C. R. Dominance and prestige: dual strategies for navigating social hierarchies. in *Advances in Experimental Social Psychology* (eds. Olson, J. M. & Zanna, M. P.) vol. 54 129–180 (Academic Press, 2016).
94. Zaidi, A. A. *et al.* Facial masculinity does not appear to be a condition-dependent male ornament and does not reflect MHC heterozygosity in humans. *Proc. Natl. Acad. Sci.* **116**, 1633–1638 (2019).
95. Kordsmeyer, T. L. & Penke, L. Effects of male testosterone and its interaction with cortisol on self- and observer-rated personality states in a competitive mating context. *J. Res. Personal.* **78**, 76–92 (2019).
96. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).
97. Mahrholz, G., Belin, P. & McAleer, P. Judgements of a speaker's personality are correlated across differing content and stimulus type. *PLoS ONE* **13**, e0204991 (2018).
98. Boersma, P. & Weenink, D. *Praat: doing phonetics by computer*. (2018).
99. Idris, F. P., Wan, Y., Zhang, X. & Punyadeera, C. Within-day baseline variation in salivary biomarkers in healthy men. *OMICs J. Integr. Biol.* **21**, 74–80 (2017).
100. Schultheiss, O. C. & Stanton, S. J. Assessment of salivary hormones. in *Methods in social neuroscience* 17–44 (eds. Harmon-Jones, E. & Beer J.S.) (Guilford Press, New York, NY 2009).
101. Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T. & Schwartz, E. B. The “trouble” with salivary testosterone. *Psychoneuroendocrinology* **29**, 1229–1240 (2004).
102. Mehta, P. H., Welker, K. M., Zilioli, S. & Carré, J. M. Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology* **56**, 88–99 (2015).
103. Kordsmeyer, T. L., Stern, J. & Penke, L. 3D anthropometric assessment and perception of male body morphology in relation to physical strength. *Am. J. Hum. Biol.* **31**, e23276 (2019).

104. Price, M. E., Dunn, J., Hopkins, S. & Kang, J. Anthropometric correlates of human anger. *Evol. Hum. Behav.* **33**, 174–181 (2012).
105. Sell, A. *et al.* 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 (2009).
106. Team, R. C. R: *A Language and Environment for Statistical Computing.* (Vienna, Austria, 2013).
107. Wickham, H. *et al.* *Ggplot2: create elegant data visualisations using the grammar of graphics.* (2020).
108. Hittner, J. B., May, K. & Silver, N. C. A Monte Carlo evaluation of tests for comparing dependent correlations. *J. Gen. Psychol.* **130**, 149–168 (2003).
109. Lenth, R. *rsm: Response-Surface Analysis.* (2020).
110. Viechtbauer, W. *metafor: Meta-Analysis Package for R.* (2020).
111. Hayes, A. F. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (The Guilford Press, New York, 2013).
112. Raiche, G. & Magis, D. *nFactors: Parallel analysis and other non graphical solutions to the cattell scree test.* (2020).
113. Revelle, W. *psych: Procedures for Psychological, Psychometric, and Personality Research.* (2020).
114. Zakzanis, K. K. Statistics to tell the truth, the whole truth, and nothing but the truth Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Arch. Clin. Neuropsychol.* **16**, 653–667 (2001).
115. Brunswik, E. *The Conceptual Framework of Psychology* (University of Chicago Press, Chicago, 1952).
116. Zhang, J., Zheng, L., Zhang, S., Xu, W. & Zheng, Y. Vocal characteristics predict infidelity intention and relationship commitment in men but not in women. *Personal. Individ. Differ.* **168**, 110389 (2021).
117. Li, Y. *et al.* Women's preference for masculine traits is disrupted by images of male-on-female aggression. *PLoS ONE* **9**, e110497 (2014).
118. Borras-Guevara, M. L., Batres, C. & Perrett, D. I. Aggressor or protector? Experiences and perceptions of violence predict preferences for masculinity. *Evol. Hum. Behav.* **38**, 481–489 (2017).
119. Dejonckere, P. H. *et al.* Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Rev. Laryngol. Otol. Rhinol.* **117**, 219–224 (1996).
120. Michaelis, D., Fröhlich, M. & Strube, H. W. Selection and combination of acoustic features for the description of pathologic voices. *J. Acoust. Soc. Am.* **103**, 1628–1639 (1998).
121. Virginia, W., James, F. & Richard, C. Acoustic prediction of severity in commonly occurring voice problems J. *Speech Lang. Hear. Res.* **38**, 273–279 (1995).
122. Puts, D. A., Doll, L. M. & Hill, A. K. Sexual selection on human voices. in *Evolutionary Perspectives on Human Sexual Psychology and Behavior* (eds. Weekes-Shackelford, V. A. & Shackelford, T. K.) 69–86 (Springer, Berlin, 2014).
123. Cheng, J. T., Tracy, J. L., Ho, S. & Henrich, J. Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans. *J. Exp. Psychol. Gen.* **145**, 536–547 (2016).
124. Altman, D. G. & Royston, P. The cost of dichotomising continuous variables. *BMJ* **332**, 1080 (2006).
125. Leary, C. J. & Knapp, R. The stress of elaborate male traits: integrating glucocorticoids with androgen-based models of sexual selection. *Anim. Behav.* **89**, 85–92 (2014).
126. Sapolsky, R. M., Romero, L. M. & Munck, A. U. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocr. Rev.* **21**, 55–89 (2000).
127. Roberts, M. L., Buchanan, K. L., Hasselquist, D. & Evans, M. R. Effects of testosterone and corticosterone on immunocompetence in the zebra finch. *Horm. Behav.* **51**, 126–134 (2007).
128. Kandrik, M. *et al.* Does the interaction between cortisol and testosterone predict men's facial attractiveness?. *Adapt. Hum. Behav. Physiol.* **3**, 275–281 (2017).
129. Dekkers, T. J. *et al.* A meta-analytical evaluation of the dual-hormone hypothesis: does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy?. *Neurosci. Biobehav. Rev.* **96**, 250–271 (2019).
130. Grebe, N. M. *et al.* Testosterone, cortisol, and status-striving personality features: a review and empirical evaluation of the Dual Hormone hypothesis. *Horm. Behav.* **109**, 25–37 (2019).
131. Feinberg, D. R., Jones, B. C. & Armstrong, M. M. Sensory exploitation, sexual dimorphism, and human voice pitch. *Trends Ecol. Evol.* **33**, 901–903 (2018).
132. Dawkins, M. S. & Guilford, T. The corruption of honest signalling. *Anim. Behav.* **41**, 865–873 (1991).
133. Andersson, M. Why are there so many threat displays?. *J. Theor. Biol.* **86**, 773–781 (1980).
134. Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F. & Winter, A. S. Vocal indices of stress: a review. *J. Voice* **27**(390), e21-390.e29 (2013).
135. Zhang, J. *et al.* Elevated recognition accuracy for low-pitched male voices in men with higher threat potential: Further evidence for the retaliation-cost model in humans. *Evol. Hum. Behav.* (2020).136.Zhang, J. & Reid, S. A. Aggression in young men high in threat potential increases after hearing low-pitched male voices: two tests of the retaliation-cost model. *Evol. Hum. Behav.* **38**, 513–521 (2020).
136. Zhang, J. & Reid, S. A. Aggression in young men high in threat potential increases after hearing low-pitched male voices: two tests of the retaliation-cost model. *Evol. Hum. Behav.* **38**, 513–521 (2017).
137. Lakens, D. One-sided tests: Efficient and underused. *The 20% Statistician* <https://daniellakens.blogspot.com/2016/03/one-sided-tests-efficient-and-underused.html> (2016).

Author contributions

C.S., L.P., and D.A.P. conceived of the presented idea. C.S., L.P., and D.A.P. designed and planned the experiments. C.S., T.L.K., and D.A.P. planned the data collection and collected the data. R.A.C. programmed the rating study. C.S. and T.A. analyzed the data. C.S. took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-77940-z>.

Correspondence and requests for materials should be addressed to D.A.P. or L.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020, corrected publication 2021