

## The Geneva Faces and Voices (GEFAV) database

Camille Ferdenzi · Sylvain Delplanque ·  
Ines Mehu-Blantar · Katharin Mylena Da Paz Cabral ·  
Marisa Domingos Felicio · David Sander

Published online: 16 December 2014  
© Psychonomic Society, Inc. 2014

**Abstract** Although many facial and vocal databases are available for research, very few of them have controlled the range of attractiveness of the stimuli that they offer. To fill this gap, we created the GENEVA Faces and Voices (GEFAV) database, providing standardized faces (static and dynamic neutral, smiling) and voices (speaking sentences, vowels) of young European adults. A total of 61 women and 50 men 18–35 years old agreed to be part of the GEFAV stimuli, and two rating studies involving 285 participants provided evaluations of the facial and vocal samples. The final set of stimuli was satisfactory in terms of attractiveness range (wide and rather symmetrical distribution over the attractiveness continuum) and the reliability of the ratings (high consistency between the two rating studies, high interrater agreement in the final rating study). Moreover, the database showed an adequate validity, since a series of findings described by earlier research on human attractiveness were confirmed—namely, that facial and vocal attractiveness are predicted by femininity and health in women, and by masculinity, dominance, and trustworthiness in men. In future studies, the GEFAV stimuli may be used intact or transformed, individually or in multimodal combinations, to investigate a wide range of mechanisms, such as the behavioral, neuropsychological, and neurophysiological processes involved in social cognition.

**Electronic supplementary material** The online version of this article (doi:10.3758/s13428-014-0545-0) contains supplementary material, which is available to authorized users.

C. Ferdenzi · S. Delplanque (✉) · I. Mehu-Blantar ·  
K. M. Da Paz Cabral · M. Domingos Felicio · D. Sander  
Swiss Center for Affective Sciences, University of Geneva, Case  
Postale 60, CH-1211 Geneva 20, Switzerland  
e-mail: Sylvain.Delplanque@unige.ch

C. Ferdenzi · S. Delplanque · I. Mehu-Blantar ·  
K. M. Da Paz Cabral · M. Domingos Felicio · D. Sander  
Laboratory for the Study of Emotion Elicitation and Expression (E3  
Lab), Department of Psychology, FPSE, University of Geneva,  
Geneva, Switzerland

**Keywords** Faces · Voices · Database · Attractiveness · Social cognition · Mate choice

The attractiveness of faces, voices, bodies, and other physical attributes, such as body odors, clearly influences social interactions. Judgments of and behaviors toward others in actual interactions are positively influenced by attractiveness (Dion, Berscheid, & Walster, 1972; Langlois et al., 2000). In mating contexts, attractiveness resides in perceptual characteristics that are considered honest indicators of mate quality, health and reproductive value (Gangestad & Scheyd, 2005; Perrett, 2010; Roberts & Little, 2008). For example, facial symmetry, which cues developmental stability (i.e., robustness to challenging environmental conditions during ontogeny), is a rather consensual criterion for facial attractiveness in humans (Grammer & Thornhill, 1994; Perrett et al., 1999; Rhodes, 2006). Furthermore, men who are able to pay the costs of high levels of testosterone by resisting its immunosuppressive properties develop masculine phenotypes such as masculine facial features or deeper voices that tend to be found attractive to women (Feinberg, Jones, Little, Burt, & Perrett, 2005; Folstad & Karter, 1992; Penton-Voak & Chen, 2004). Despite beauty criteria indisputably having selective advantages, perception of them by conspecifics varies according to historical (Ricciardelli & Williams, 2012; Senticilles & Callahan, 2012), cultural (DeBruine, Jones, Crawford, Welling, & Little, 2010), and individual factors (e.g., own mate quality, relationship context, phase of the menstrual cycle, or partnership status; see Little, Jones, & DeBruine, 2011, for a review).

Attractiveness studies in evolutionary psychology, but also studies on voice and face perception in other disciplines (e.g., social and affective sciences), often use the presentation of stimuli such as photographs of faces and audio recordings of voices. Constructing these stimulus sets is costly and challenging, in terms of participant recruitment, data collection,

standardization of recording conditions and so on. Furthermore, ethical restrictions regarding use of the participants' image or voice may place constraints on their reuse by the original researchers or on sharing these resources with other interested researchers. A straightforward solution to all of these issues would be to use stimuli from a database.

To our knowledge, databases of voice samples exist for aging voices (Kelly, Drygajlo, & Harte, 2013), pathological speech/voices (e.g., Disordered Voice Database, Kay Elemetrics Corp., Lincoln Park, NJ, USA), and emotional voices (Belin, Fillion-Bilodeau, & Gosselin, 2008; Burkhardt, Paeschke, Rolfes, Sendlmeier, & Weiss, 2005; Petta, Pelachaud, & Cowie, 2011). Most facial databases were designed for the study of face recognition (Gross, 2005; [www.face-rec.org/databases](http://www.face-rec.org/databases)) and/or facial expression recognition, such as the Pictures of Facial Affect (Ekman & Friesen, 1976), the Karolinska Directed Emotional Faces (Lundqvist, Flykt, & Öhman, 1998), the PAL (Minear & Park, 2004), the Yale Face Database (Belhumeur, Hespanha, & Kriegman, 1997), the PICS (University of Stirling, <http://pics.psych.stir.ac.uk>), and the FACES (Max Planck Institute for Human Development in Berlin, <http://faces.mpib-berlin.mpg.de>). Several multimodal databases including faces and voices have also been developed for emotion research (Bänziger, Mortillaro, & Scherer, 2012; Petta et al., 2011).

Unfortunately, each database has its limits, and researchers willing to use stimuli contrasted in attractiveness have to deal with images that are not sufficiently standardized (in terms of background, clothing, hair positioning, or facial expression), are slightly outdated, include various age groups and ethnicities (thus offering only small homogeneous subsamples), or have restricted access. Importantly, most databases do not provide stimuli that are distributed over a useful range of attractiveness. Among the very few databases of faces with controlled levels of attractiveness, one created by Todorov and colleagues consists of 25 computer-generated faces (not real human faces) varying on seven levels of attractiveness, and is freely available after completion of an agreement form (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013). Another database (the Beautycheck project; Braun, Gruendl, Marberger, & Scherber, 2001) provides composite pictures derived from three or more original faces among neutral standardized pictures of 96 young adults (no original face available).

In this context, our aim was to create a new database, the GENEVA Faces and Voices (GEFAV) database, to be used both in attractiveness research and in much wider areas of face and voice perception. The GEFAV fits the following criteria: original stimuli from real individuals, different types of stimuli from the same individuals (visual static and dynamic; auditory), homogeneous sample (young adults, men and women, European type), stimuli collected in highly standardized conditions, and free access for academic nonprofit research.

Compared with existing databases, the main contribution of the GEFAV is that particular care was taken to have stimuli covering a sufficient range of attractiveness, based on evaluations by a population with the same characteristics as the stimuli. This was achieved by a two-step procedure: Study 1 allowed for stimulus collection and the ratings of 73 GEFAV participants, and Study 2 was conducted to populate the higher tail of the attractiveness distribution and to collect perceptual judgments of the final 111 GEFAV participants on several social dimensions, including attractiveness. This report introduces the GEFAV database and shows its quality through (1) the appropriateness of the attractiveness range; (2) its reliability, represented by rater consistency within and between studies; and (3) its validity, confirmed by attractiveness predictions by classical factors such as face symmetry (Grammer & Thornhill, 1994; Perrett et al., 1999), digit ratio (Bogaert, Fawcett, & Jamieson, 2009; Ferdenzi, Lemaitre, Leongomez, & Roberts, 2011), voice frequency (Collins, 2000; Collins & Missing, 2003), and voice distance from the mean frequency of the sample (Bestelmeyer et al., 2012; Bruckert et al., 2010).

## Material and method

### Stimuli

The GEFAV database comprises facial and vocal stimuli from 111 participants, hereafter called "stimulus participants" (61 women, 50 men), 18–35 years old (mean  $\pm$  SD = 22.9  $\pm$  4.3 years). Among women, 35 reported using hormonal contraceptives and 26 were not. The latter were recorded during the fertile phase of their menstrual cycle (based on the estimation method detailed in Ferdenzi, Schaal, & Roberts, 2009). Eligible stimulus participants were 18 to 35 years of age, of European origin, French native speakers, nonsmokers, and did not have any form of speech impediment or visible ornaments or artifacts (e.g., piercings or braces). They were recruited by posters and Web announcements (Study 1, untargeted recruitment) and by word of mouth or via models' webpages (Study 2, targeted at higher levels of physical attractiveness). They received cash or course credits for their participation.

The Committee on Research Ethics of the Faculty of Psychology and Education Sciences at the University of Geneva approved this study and the constitution of the database. An information sheet (Appendix A) explained to the participants that they were going to be videotaped and audio recorded, that their face and voice were going to be rated by others on several criteria, that their face and voice were going to be used in later studies if they agree to take part in the GEFAV database, and that their name will never be associated to their stimuli when communicating the results/the stimuli to third persons. After reading this document, the participants signed a consent form

(Appendix B), in which they indicated whether they agreed with each of the following points separately: (1) to have their face videotaped and their voice recorded, and to have their face and voice subsequently evaluated by other participants, and (2) to take part with their face and voice in the constitution of a permanent database that would subsequently be used in other research projects. All participants included in the GEFAV gave written consent to these two points. Participants who agreed only to the first point were not included in the GEFAV and their data served a secondary purpose of this study not presented in this article. Later on, as a fulfillment of participants' rights to have more precise information about how their stimuli could be used, we explained to them by e-mail that (1) the database would be made available for free to other researchers from other institutions than the University of Geneva, upon request and for academic research only; (2) that these researchers, by signing the GEFAV Confidentiality Agreement and Terms of Use (Appendix C), would commit to using the stimuli of the database for academic research only and not to communicate the database to others; and (3) that their face and voice could be presented in other studies in their original form or in a modified form (e.g., through masculinization of the voice/face, combination with other faces/voices, or transformation of the face into a computer-generated face). All participants from whom we received an acknowledgment of receipt were kept in the database ( $N = 111$ ), and we removed participants who replied that they did not agree with these terms ( $N = 2$ ) or had an invalid e-mail address ( $N = 1$ ). The GEFAV stimuli can be considered as personal data because they allow direct or indirect identification of the person. Therefore, we made sure that our processing of these data is in compliance with the European legislation (European Directive 95/46/EC) and the Swiss Federal Act on Data Protection. Only researchers from European countries and from countries benefiting from a decision of the European Commission regarding the adequacy of protection of personal data in third-party countries [based on Article 25(6) of directive 95/46/EC], such as Switzerland (reference of the decision: 2000/518/EC), are allowed to have access to the GEFAV. Researcher requests for access to the database will be examined case-by-case for this purpose. Conditions of data transfer and content of the database are subject to change according to possible modifications of the legislation in the future.

The stimulus participants were simultaneously video- and audio-recorded with a Panasonic AG-HPX171 video camera and a Neumann BCM 104 condenser studio microphone, during a single session. Men were asked to have their face shaved and women were instructed not to wear any makeup. Despite these explicit instructions, several men had stubble on the day of testing, but we included them in spite of this and let the database users decide whether they wish to use stimuli with facial hair or not. Also, a few women did not comply with the makeup instructions and were asked to

remove it with facial wipes that we provided. In spite of this procedure, some traces of eye pencil can be seen on some of the participants. Before the recording took place, participants were asked to remove all accessories, to put on a black top, to make their face free from hair, and to sit in front of a black background. Standardized conditions of light, sound, material, and participant's placement were carefully applied. Participants were asked (1) to look straight at the camera with a neutral expression, (2) to say *Bonjour. Il est deux heures moins dix* ("Hello. It's ten to two") toward the camera with a neutral expression, (3) to freely express their emotion as they looked at the camera while giggles were played aloud in the room (the aim being to elicit a smile), and (4) to pronounce the six vowels / $\epsilon$ /, / $i$ /, / $a$ /, / $o$ /, / $u$ /, and / $y$ / (from the International Phonetic Alphabet). This allowed us obtaining three color facial stimuli: a static neutral picture and a static smiling picture, both extracted from the video, and a short video sequence showing the face saying the sentence. Using Psychomorph (Tiddeman, Burt, & Perrett, 2001) and Adobe Creative Suite, faces were equalized in width, and masked to remove hair and retain only the face, ears and neck. In addition, two vocal stimuli were extracted: the sentence and the vowels / $i$ /, / $a$ /, and / $o$ / (from the middle of the vowel series, to limit intonation variations; Collins, 2000). A 2-s duration was set for the vowel series with Praat version 5.2 (Boersma & Weenink, 2011), by shortening or lengthening the silent periods between vowels (not the vowels themselves; see Ferdenzi et al., 2013). Sound intensity was normalized with MATLAB version 7.12 by matching the average absolute amplitude of all recordings. Examples of the different types of samples are provided as Supplementary Samples 1–5, and further details about stimulus collection and preparation are available in the Supplementary Method.

Additional stimuli and measures were used in this study. First, the lengths of index and ring fingers of the stimulus participants were measured, to allow for computing the second-to-fourth digit ratio (2D:4D) of the right hand. Second, body mass index (BMI) was estimated on the basis of participants' reports of their height and weight. Third, an index of facial symmetry was computed with Psychomorph, on the basis of the distances between seven bilateral points of the face (Scheib, Gangestad, & Thornhill, 1999), and fundamental frequency (F0) of the voices was measured with Praat. These four parameters were measured because they are likely to be related with attractiveness. Finally, five male and five female composite faces were constituted and used in Study 2 (Supplementary Fig. 1): These were face averages of the stimulus participants with the lowest, highest, and intermediate levels of attractiveness in Study 1, and with the lowest and highest finger-length ratios 2D:4D. More details about these additional stimuli and measures are available in the Supplementary Method section.

## Ratings

The rating sessions involved 98 participants in Study 1 (57 women and 41 men,  $23.3 \pm 3.8$  years old) and 187 other participants in Study 2 (91 women, 96 men;  $22.1 \pm 3.2$  years old), hereafter called “raters.” Some of the raters in Study 1 ( $N = 61$ ) were also GEFAV stimulus participants. To take part in the rating sessions, participants had to be 18–35 years of age, of European origin, heterosexual, and French native speakers. Normal vision and hearing was also required. Raters in both studies received cash or course credits as compensation to their participation.

The rating procedures in Studies 1 and 2 were mostly similar (for details, see the Supplementary Method). After providing written informed consent, raters were positioned in front of a computer in a testing room with up to ten participants simultaneously. The stimuli were a subsample of the GEFAV stimuli in Study 1 (from  $N = 73$  participants) and the total set in Study 2 (from  $N = 111$ , including 38 new participants from the targeted recruitment). Faces were presented on a screen and voices through headphones. At the beginning of the session, the raters had a quick glimpse at the faces and voices they were about to evaluate, so that they have an idea of the attractiveness range. Blocks of the same stimulus type were presented, separated by minimum 30-s breaks. Only opposite-sex stimuli were presented in Study 1, whereas in Study 2 raters evaluated male and female stimuli. In Study 2, average faces (see the Stimuli section) were also presented at the end of the session. Raters were asked to skip the ratings of stimulus participants they knew. In Study 1, stimuli were rated for attractiveness short-term (i.e., for a casual relationship), attractiveness long-term (i.e., for a lasting relationship), beauty, masculinity/femininity, and trustworthiness. In Study 2, raters evaluated attractiveness, masculinity/femininity, dominance, health, and trustworthiness, as well as perceived symmetry (only for static and dynamic neutral faces and for averaged faces) and smile authenticity (only for static smiling faces). We defined attractiveness and explained it to the participants as the propensity of the stimulus person to elicit attraction to them personally, or in general when it comes to evaluate same-sex stimuli (e.g., heterosexual males are able to project how attractive another heterosexual male would be to women). Note that the ratings/methods used in Studies 1 and 2 slightly differ due to the fact that other research questions were addressed concurrently. Especially, the distinction between beauty and attractiveness will be presented elsewhere (Ferdenzi et al., 2014). Visual analogue scales were used, and detailed explanation of each scale was provided at the beginning and during the session. For comparison purpose with Study 2, and because short-term and long-term attractiveness ratings were highly correlated ( $r_s = .94$  to  $.98$ , depending on the stimulus type), these two variables were averaged to provide a single attractiveness variable.

## Results

### Level and range of attractiveness

Analyses were conducted on mean ratings obtained by averaging raters' responses for each GEFAV stimulus ( $N = 16$  to 32 opposite-sex raters in Study 1, and  $N = 20$  to 23 opposite- and same-sex raters in Study 2). The means and ranges of ratings in Studies 1 and 2 are given in Table 1 for attractiveness, and in Supplementary Table 1 for the other dimensions. First, we tested whether targeted recruitment improved the attractiveness range of the face database. In Study 2, the new faces obtained through targeted recruitment ( $N = 38$ ) were more attractive than the faces obtained through nontargeted recruitment ( $N = 73$ ) [static:  $47.6 \pm 13.4$  vs.  $37.7 \pm 13.3$ ,  $t(109) = 3.71$ ,  $p = .0003$ ; dynamic:  $49.2 \pm 14.7$  vs.  $41.1 \pm 12.1$ ,  $t(107) = 3.11$ ,  $p = .0024$ ]. In addition, although the attractiveness distribution was positively skewed in Study 1 (skewness = 1.08 for static and 0.71 for dynamic faces), attractiveness scores were more evenly distributed in Study 2 (skewness = 0.26 and 0.16) and did not differ from normal distributions (Kolmogorov–Smirnov tests; Fig. 1). Note that the voices obtained through the two types of (face-directed) recruitment did not differ in attractiveness [sentence:  $t(106) = 1.02$ ,  $p = .309$ ; vowels:  $t(109) = 0.31$ ,  $p = .756$ ].

To briefly describe the GEFAV attractiveness data, we conducted two repeated measures analyses of variance (ANOVAs) on the attractiveness ratings from Study 2, with Sex of the Stimulus Participant as a between-subjects factor and Rater Sex and Stimulus Type (faces: static neutral vs. static smiling vs. dynamic neutral; voices: sentence vs. vowels) as within-subjects factors. Only first-order interactions were considered. For faces, we only found a significant main effect of stimulus type [ $F(1.94, 207.65) = 9.90$ ,  $p = .0001$ , Greenhouse–Geisser corrected]: videos and static smiling faces were more attractive than static neutral faces (post-hoc Tukey HSD test). Note that the attractiveness ratings of these three types of stimuli were highly correlated (neutral–smiling,  $r = .85$ ; neutral–video,  $r = .87$ ; smiling–video,  $r = .81$ ; all  $p_s < .0001$ ). For voices, the sentences and vowels did not differ significantly (and were correlated:  $r = .47$ ,  $p < .0001$ ), but there was a significant interaction between rater sex and participant sex [ $F(1, 106) = 69.20$ ,  $p < .0001$ ] due to the fact that ratings of opposite-sex voices were higher than ratings of same-sex voices (post-hoc Tukey HSD test).

### Reliability

First, using Cronbach's alpha coefficients, we tested whether the  $N = 73$  stimuli used in both Studies 1 and 2 received consistent judgments by the two different groups of raters. We found good interrater agreement ( $\alpha > .70$ ; Kline, 1993) for the static ( $\alpha = .88$ ) and dynamic ( $\alpha = .83$ ) neutral faces, for the

**Table 1** Means and ranges (minimum–maximum) of attractiveness ratings on a 0–100 scale for static neutral, dynamic neutral, and static smiling faces, and for sentence and vowels

|                       | Opposite-Sex Ratings |                                   |                                   |                                   | Same-Sex Ratings                  |                                   |
|-----------------------|----------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
|                       | Female Stimuli       |                                   | Male Stimuli                      |                                   | Female Stimuli                    | Male Stimuli                      |
|                       | S1 ( <i>N</i> = 36)  | S2 ( <i>N</i> = 61 <sup>a</sup> ) | S1 ( <i>N</i> = 37 <sup>b</sup> ) | S2 ( <i>N</i> = 50 <sup>b</sup> ) | S2 ( <i>N</i> = 61 <sup>a</sup> ) | S2 ( <i>N</i> = 50 <sup>b</sup> ) |
| Face: Static neutral  | 32.5 (10–76)         | 41.5 (16–76)                      | 26.8 (6–72)                       | 39.7 (14–81)                      | 43.7 (15–71)                      | 38.6 (7–71)                       |
| Face: Dynamic neutral | 37.0 (11–86)         | 44.1 (11–78)                      | 31.4 (8–69)                       | 44.2 (11–75)                      | 45.7 (10–75)                      | 41.3 (10–78)                      |
| Face: Static smiling  | n/a                  | 42.8 (15–78)                      | n/a                               | 41.8 (11–79)                      | 46.6 (11–73)                      | 42.4 (10–70)                      |
| Voice: Sentence       | 51.1 (28–79)         | 58.0 (25–86)                      | 47.6 (13–74)                      | 51.9 (24–76)                      | 51.0 (20–73)                      | 45.5 (23–75)                      |
| Voice: Vowels         | 51.2 (25–67)         | 57.6 (31–80)                      | 48.2 (13–74)                      | 54.3 (20–83)                      | 50.7 (21–71)                      | 45.5 (17–69)                      |

Only opposite-sex ratings were collected in Study 1 (S1, involving 73 stimuli collected through nontargeted recruitment), whereas both same- and opposite-sex ratings were collected in Study 2 (S2, involving an additional 38 stimuli collected through recruitment targeted at attractive individuals). <sup>a</sup>Minus 1 for Voice: Sentence (missing stimulus). <sup>b</sup>Minus 2 for Face: Dynamic neutral and Voice: Sentence (missing stimuli).

vowels ( $\alpha = .83$ ) and the sentences ( $\alpha = .74$ ; this value is slightly lower probably because the sentence excerpts slightly differed between Studies 1 and 2; see the Supplementary Method).

Second, we tested whether the Study 2 attractiveness evaluations of the average faces (see the **Stimuli** section and Supplementary Fig. 1) were consistent with the Study 1 ratings. Greenhouse–Geisser-corrected repeated measures ANOVAs on opposite-sex attractiveness ratings showed a significant effect of attractiveness [male averages:  $F(1.89, 164.40) = 123.59, p < .0001$ ; female averages:  $F(1.99, 189.50) = 90.87, p < .0001$ ]. Indeed, for both sexes the face averages made of the least attractive stimuli of Study 1 were less attractive in Study 2 than the averages made of the most attractive stimuli of Study 1 (post-hoc Tukey HSD tests; see Supplementary Fig. 2a).

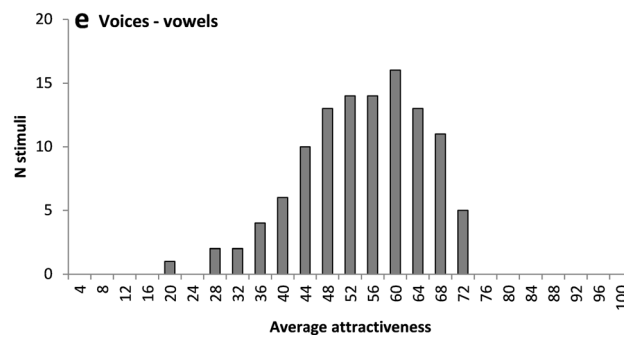
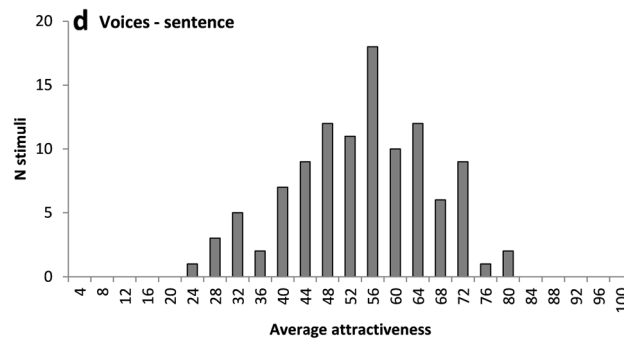
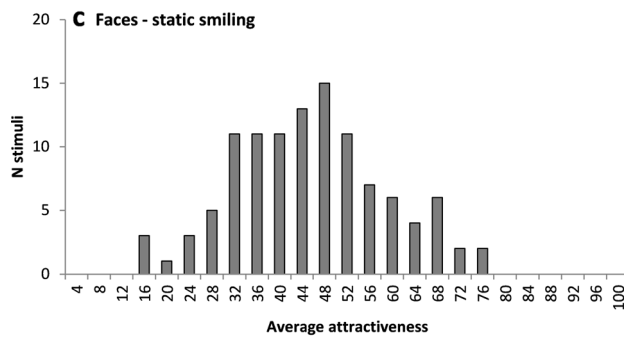
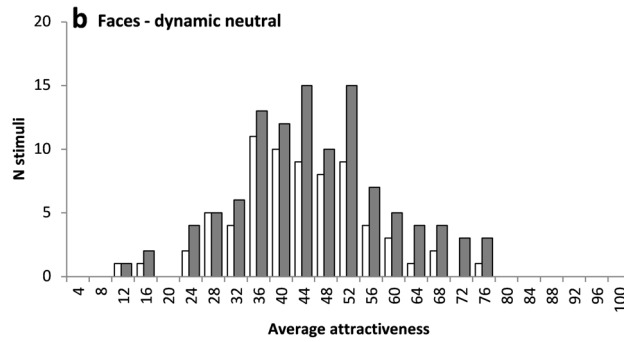
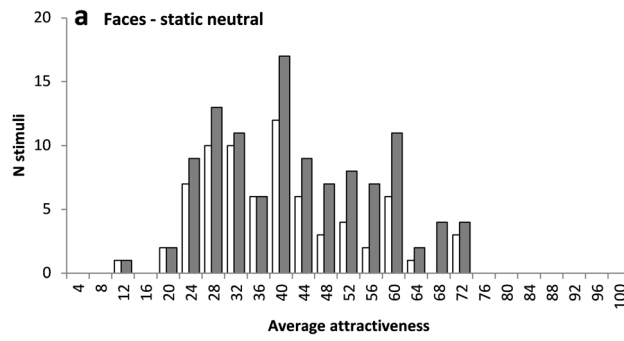
A third test of reliability consisted in evaluating the interrater agreement in Study 2 for each stimulus. Cronbach's alphas were satisfactory:  $\alpha = .91$  for static neutral faces,  $\alpha = .90$  for static smiling faces,  $\alpha = .88$  for dynamic neutral faces,  $\alpha = .87$  for sentences, and  $\alpha = .83$  for vowels (averages of the nine  $\alpha$  coefficients for the nine groups of raters evaluating nine different sets of stimuli; see the Supplementary Method).

### Validity

In a first series of analyses on Study 2 ratings, we tested separately whether several factors, known to be significant determinants of face and voice attractiveness, predicted attractiveness also in the GEFAV. As expected, simple regressions revealed that face symmetry predicted attractiveness [static neutral faces:  $r^2 = .15, F(1, 109) = 19.04, p < .0001$ ; dynamic neutral faces:  $r^2 = .29, F(1, 107) = 43.26, p < .0001$ ; see Supplementary Fig. 3]. Similar analyses on voice fundamental frequency showed that, consistent with the literature, female

voices with higher frequencies were more attractive [sentence:  $r^2 = .07, F(1, 58) = 4.10, p = .048$ ; vowels:  $r^2 = .09, F(1, 59) = 5.98, p = .017$ ] and male voices with lower frequencies were more attractive [sentence:  $r^2 = .09, F(1, 46) = 4.29, p = .044$ ; vowels:  $r^2 = .37, F(1, 48) = 27.88, p < .0001$ ; see Supplementary Fig. 4]. In addition, voices' distance to the mean—namely, the distance to the sex-specific average fundamental frequency computed from our sample—negatively predicted voice attractiveness [sentence:  $r^2 = .05, F(1, 106) = 5.73, p = .018$ ; vowels:  $r^2 = .08, F(1, 109) = 9.60, p = .0025$ ; see Supplementary Fig. 5]. Finally, a repeated measures ANOVA was conducted on the attractiveness of the faces averaged according to 2D:4D. The face average made from the stimulus women with the most feminine 2D:4D was more attractive than the one made from the stimulus women with the most masculine 2D:4D [ $F(1, 184) = 43.45, p < .0001$ ]. The reverse was found for male faces [ $F(1, 184) = 11.35, p = .0009$ ], with the face average made from the stimulus men with the most masculine 2D:4D being more attractive (see Supplementary Fig. 2b). Note that in this paragraph, the same conclusions were found when only opposite-sex ratings were taken into account.

Another way to test the validity of the GEFAV was to investigate which combination of factors, among the ones that we measured (see Table 2), best predicted attractiveness. We ran a model comparison procedure on the opposite-sex attractiveness ratings using R 3.0.1 (R Development Core Team, 2011). For each stimulus sex and type, we used the “dredge” function of the MuMIn R package (Barton, 2013) to run all possible linear regressions involving all possible combinations of factors. The best models were the ones with the lowest Akaike information criterion (AIC), and when two or more models had an AIC difference less than 2, we retained the simplest model in order to be parsimonious (see Table 2 for the best models of Study 2, and Supplementary Table 2 for



◀ **Fig. 1** Distributions of attractiveness scores (0–100) averaged by GEFAV participant, for (a) static neutral faces, (b) dynamic neutral faces, (c) static smiling faces, (d) sentences, and (e) vowels. For the neutral faces, the distribution of the attractiveness scores before targeted recruitment (Study 1) appears in white (Study 2 results are the gray bars).

Study 1). The two studies showed consistent differences between male and female stimuli. Women whose face and voice were perceived as being healthier and more feminine were more attractive, and men whose face and voice were perceived as being more dominant/masculine (as perceived and measured via voice frequency) were more attractive. Trustworthiness also predicted attractiveness, mostly in men, independently of relationship context, and—to a lesser extent—in women, but only in the context of long-term relationships (Supplementary Table 2). Face symmetry and 2D:4D were not significant predictors in these models, whereas they were when taken individually, suggesting that these measures explain a less significant part of attractiveness variance than the other factors included in the models.

**Table 2** Best-fitting linear models explaining attractiveness (Attrac) in Study 2 as a function of various parameters, obtained after testing all possible models and parameter combinations (opposite-sex ratings of 111 individuals)

| Stimulus Type        | Best Linear Model                    | AIC    | $r^2$ |
|----------------------|--------------------------------------|--------|-------|
| <b>Female Faces</b>  |                                      |        |       |
| Static neutral       | Attrac = Health + Fem                | 432.57 | .70   |
| Static smiling       | Attrac = Health + Fem + Domin        | 433.02 | .68   |
| Dynamic neutral      | Attrac = Health + Fem                | 423.25 | .72   |
| <b>Female Voices</b> |                                      |        |       |
| Sentence             | Attrac = Health + Fem + Trust + F0   | 418.15 | .73   |
| Vowels               | Attrac = Health + Fem + Trust        | 423.95 | .52   |
| <b>Male Faces</b>    |                                      |        |       |
| Static neutral       | Attrac = Domin + Trust               | 332.63 | .71   |
| Static smiling       | Attrac = Domin + Trust               | 317.76 | .77   |
| Dynamic neutral      | Attrac = Domin + Trust + Symm        | 339.33 | .63   |
| <b>Male Voices</b>   |                                      |        |       |
| Sentence             | Attrac = Domin + Trust               | 325.73 | .62   |
| Vowels               | Attrac = Domin + Trust + F0 + Height | 276.36 | .88   |

The parameters investigated were participant's age (not in the table); body mass index (BMI); body height (for voices only; since the sizes of the faces were equalized, this variable was not relevant for faces); second-to-fourth digit ratio of the right hand (2D:4D); dominance (Domin); femininity (Fem); masculinity (not in the table); health; trustworthiness (trust); voice fundamental frequency (F0; for voices only); face symmetry, both rated (Symm; for faces only) and measured (not in the table; for faces only); and smile authenticity (not in the table; for smiling faces only). AIC = Akaike's information criterion. All models were significant at  $p < .001$ .

## Discussion

This article introduces a new database of social stimuli, the GEFAV database. Originating from 61 women and 50 men of European type, 18–35 years of age (see Supplementary Fig. 6 for an overview of the faces), the GEFAV stimuli are meant to be used in studies investigating human perception of faces and voices varying on attractiveness. As compared with other existing databases, the GEFAV stands out for the controlled range of attractiveness that it offers, for the diversity of the stimuli (facial and vocal, static and dynamic, smiling and neutral), and for its free access under certain conditions.

The GEFAV database successfully met three quality criteria. First, in the final set of stimuli, the range of attractiveness was satisfactory—namely, 7–81 for faces and 17–86 for voices, on a 0–100 scale—and the attractiveness distribution across facial stimuli became more symmetrical after a targeted recruitment phase. One of the GEFAV aims, namely offering a relatively balanced number of attractive and less attractive stimuli, was thus reached, at least for a population with the same characteristics as the stimuli. It must be kept in mind that this distribution reflects more the requirements of our database than the distribution observed in real-life. The authors of the Beautycheck database (Braun et al., 2001) also acknowledged that in reality the distribution of facial attractiveness is skewed as in our Study 1—that is, with a larger proportion of faces with low or moderate attractiveness, and fewer very attractive ones. It is therefore very important that researchers using the GEFAV keep in mind that the stimuli are appropriate for experimental studies, but may not be appropriate for correlational studies that assume an ecologically valid distribution of attractiveness.

Second, the GEFAV proved to be a reliable tool. Indeed, the same stimuli (a large subset of the GEFAV) rated by different participants in two different studies received highly consistent attractiveness ratings, in spite of the slightly different settings. In addition, averaged faces made from the most and least attractive faces, based on the first study ratings, were rated as such by the raters of the second study. Also, the interrater agreement for each stimulus type was very high in the second study, when all GEFAV stimuli were involved.

Third, we also showed the validity of the database by testing whether the face and voice attractiveness ratings could be predicted by parameters known to be influential on human attractiveness. In spite of slight variations based on whether all GEFAV stimulus participants (Study 2) or only a subsample (Study 1) were included, we found converging evidence that women perceived as more feminine and healthier were more attractive to men and that men perceived as more dominant/masculine and more trustworthy were more attractive to women. Although it cannot be excluded that the multiple ratings given by a participant may have influenced each

other (halo effect), these results are in line with previous studies on human attractiveness showing that predictors of visual and auditory attractiveness are mainly related with sexual dimorphism (Perrett et al., 1998). This was typically the case for voices, with deeper male voices and female voices with higher frequencies being rated more attractive in the present and previous research (Collins, 2000; Collins & Missing, 2003; Vukovic et al., 2011). Perceived femininity and health were significant predictors of women attractiveness because they are indicators of women's hormonal status, and thus proxies for potential reproductive fitness. Indeed, high estrogen during the fertile phase of the menstrual cycle is associated with higher femininity, attractiveness and health in the face (Law Smith et al., 2006) and higher voice femininity (Bryant & Haselton, 2009). Men's attractiveness depended on a combination of two positive and pro-social traits that are important for long-term relationships and paternal investment: assertiveness/confidence (dominance ratings) and reliability/honesty (represented by trustworthiness). How fetal sex hormones, represented by the second-to-fourth digit ratio (2D:4D), influence adult vocal and facial characteristics and attractiveness, remains to be elucidated in the future. Individuals with more sex-typical ratios were more attractive in our study as in previous ones (Bogaert et al., 2009; Ferdenzi et al., 2011; Roney & Maestripieri, 2004), a result that was, however, not replicated when other predictors were taken into account to explain attractiveness variability (Table 2). Likewise, as recurrently shown in the literature (Grammer & Thornhill, 1994; Perrett et al., 1999), rated symmetry of GEFAV faces significantly predicted their attractiveness but this result was masked when simultaneously taking into account other factors related with sexual dimorphism (Table 2). Current indices of sexual dimorphism may thus have more weight in determining attractiveness of a face, but this deserves confirmation in the future.

The GEFAV database is available upon request for scientific, nonprofit research. The stimuli are provided by the Swiss Center for Affective Sciences at the University of Geneva, Switzerland, and request for access to the database can be formulated via the following webpage: [www.affective-sciences.org/gefav](http://www.affective-sciences.org/gefav). Currently, the available stimuli are pictures of neutral and happy facial expressions (formats, .tif and .jpg; size, 576 × 720 pixels), soundless videos of the faces while pronouncing a neutral sentence (formats, .avi and .mp4; size, 720 × 576 pixels), and vocal audio recordings of a series of three vowels and of a neutral sentence (formats, .wav and .mp3). These formats will be updated as appropriate over the coming years. To facilitate and broaden the use of the GEFAV stimuli, we also provide additional useful material such as Psychomorph templates of the faces. Finally, to limit unregulated circulation and misuse of the GEFAV stimuli, each researcher with a tenured position using the database needs to register and sign an

agreement regarding the terms of use (Appendix C) before obtaining access to the database, and only researchers from European countries and from countries benefiting from a decision of the European Commission regarding adequacy of the protection of personal data are allowed to have access to the GEFAV.

Although the quality of the GEFAV has been demonstrated in this article, some limitations of this database must be acknowledged. First, the audio recording of the sentence in French may be mostly suitable for studies with French-speaking participants. The other vocal stimuli (three vowels), although sounding rather French as well, may more easily be used internationally. Second, the picture size is adequate for many usages, but may be considered as insufficient for other very specific ones, such as accurate morphometric analyses of the faces for instance. Third, although it is an asset to offer a homogeneous sample to allow researchers to have enough stimuli to choose from, researchers who are interested in various ages or ethnicities may find that the GEFAV is limited. In no case should the GEFAV be considered as a standard database or as representing a universal range of human facial attractiveness, and the development of more and more varied sets of facial and vocal stimuli is strongly encouraged. Also, the ratings presented in this article are from a specific and limited sample (participants recruited in Geneva), therefore we strongly recommend that GEFAV users have the stimuli judged to determine their attractiveness levels in the population they are interested in.

Researchers must nonetheless keep in mind that the GEFAV as it is provided on our website may constitute ground material that can be modified, transformed, and combined to get a wide range of possible new stimuli. For example, the masculinity/femininity, symmetry, skin color or appearance, and even age of the faces can be manipulated, as well as the fundamental frequency or other acoustic parameters of the voices. Multiple faces or voices can be combined by averaging them. Different multimodal combinations can be considered with our database, using simultaneous presentation of voice and face from the same person or from different individuals. The GEFAV stimuli can either be the target of the experiment, when investigating the effect of individual or contextual factors on the perception of faces or voices, or can be used as stimuli with positive or negative valence (in terms of attractiveness) to investigate specific psychological processes. In sum, researchers from disciplines as varied as social and affective sciences, evolutionary, experimental, and developmental psychology, neuroscience, and psychobiology will be able to take advantage of the GEFAV stimuli to explore verbal and nonverbal human responses to social stimuli. We hope that the GEFAV will be a useful tool to investigate further uni- and multimodal psychological mechanisms involved in attractiveness and, more generally, in social cognition.



**Author Note** C.F. is now located at the Centre de Recherche en Neurosciences de Lyon, CNRS UMR5292, INSERM U1028, Université Claude Bernard Lyon 1, 50 avenue Tony Garnier, 69366 Lyon Cedex 7, France. The authors thank Christophe Mermoud, Benoît Bediou, Sophie Jarlier, Dajana Kapusova, Julien Savary, and Olga Vorontsova for technical support; Jean-François Lemaître and Craig Roberts for useful discussions and comments on the manuscript; and three anonymous reviewers for their useful comments and suggestions. This research was financed by a grant from the Swiss National Science Foundation (100014\_130036) and was supported by the National Center of Competence in Research “Affective Sciences” (Grant No. 51NF40-104897), hosted by the University of Geneva.

### Appendix A: Information sheet given to the participants prior to participation (translated from French)

**Research title:** link between the characteristics of the face, the voice and the body odour

**Researchers in charge of the project:** Drs Camille Ferdenzi, Sylvain Delplanque and David Sander

**Dear participant, please read carefully this information sheet that explains why and how we are conducting this research. Do not hesitate to ask for additional information. If you agree to participate in the study, please complete the consent form attached.**

- **General aim of the research:**

This experiment aims to study the link between different features of the face, the voice and the body odour. Voices and faces will be used to constitute a permanent database, which will then be used in various studies on face and voice perception.

- **Procedure:**

The experimenter will videotape your face and record your voice in standard conditions in the laboratory. To do this, we request that women do not to wear any make-up and that all participants wear a black top, let their face free from hair and remove any jewellery or piercing. At the end of the session, we will also take some body measurements (finger length, and waist and hip for women / waist and shoulders for men). [...] If you agree to take part in the constitution of the database, you give us the permission to use your face and voice in future studies on face and voice perception. Your face, your voice [...] will be evaluated on various criteria (face: symmetry, masculinity / femininity, pleasantness, etc.; voice: dominance, masculinity / femininity, pleasantness, etc. [...]).

- **Protection of the data:**

Data and samples (faces, voices [...]) will be stored anonymously (you will be assigned a code) and securely. Only researchers involved in the study will have access to them. If you agree to participate in the database, we will maintain the images of your face and your voice

recordings indefinitely; otherwise, the samples will be destroyed after the publication of the results. Researchers will maintain sufficient information to be able to link your data to your name during the period of data collection, but this information will not be disclosed to others.

- **Advantages and benefits for the participants:**

Participation in this research is of interest to anyone who can read, interpret, design and conduct experiments in psychology. In addition, participation in the study is paid.

- **Disadvantages and risks for the participants:**

There are no foreseeable risks to take part in this study [...]. In case of inconvenience, please inform the experimenter.

- **Duration of the project:** 2 years

- **Duration of the experiment and breaks:**

It will take about 20 minutes (without breaks) to videotape your face and record your voice [...].

- **Access to the results of the study:**

If you wish, we will send you an overview of the results of the study (only for the group). We will publish the results of this study, but you will never be personally identified. Please contact the researchers if you have any questions about the publication of the results.

- **Contacts:**

Dr Camille Ferdenzi, Postdoctoral Researcher, Tel. 022 379 9806, Email [Camille.Ferdenzi@unige.ch](mailto:Camille.Ferdenzi@unige.ch)

Dr Sylvain Delplanque, Senior Researcher, Tel. 022 379 9826, Email [Sylvain.Delplanque@unige.ch](mailto:Sylvain.Delplanque@unige.ch)

*Note: elements of the document concerning body odour collection have been removed ([...]), since they are not relevant to the present article.*

### Appendix B: Consent form given to the participants prior to participation (translated from French)

**Research title:** link between the characteristics of the face, the voice and the body odour

**Researchers in charge of the project:** Drs Camille Ferdenzi, Sylvain Delplanque and David Sander

**Participant's consent:** Based on the information sheet, I confirm that I agree to participate in the study entitled “Link between the characteristics of the face, the voice and the body odour”, and that I agree (please tick):

- To have my face videotaped and my voice recorded, and that my face and my voice will subsequently be evaluated by other participants  yes  no
- [...]

- To take part (with my face and my voice) in the constitution of a permanent database that will subsequently be used in other research projects  yes  no
- That my data will be used for scientific purposes and that the results will be published in scientific journals or books, provided that the data will remain anonymous and no information will be given about my identity  yes  no
- That the data will be used for educational purposes (courses and seminars to train students and professionals bound by professional secrecy)  yes  no
- Would you like to receive an overview of the results of the study?  yes  no

If yes, please provide your email address in capital letters:

I have voluntarily chosen to participate in this research. I was informed that I have the right to withdraw at any time without explanations and that in such case I can request my data to be destroyed.

Participant's name, Surname, Date and Signature

**Researcher's commitment:** The information given in this form and the answers I gave to the participant describe the research accurately. I commit myself to conducting the study in accordance with the ethical standards for research involving human participants, applying the Code of Ethics for Research of the Faculty of Psychology and Educational Sciences, and the guidelines of the University of Geneva related to integrity in the field of scientific research and to the procedure to follow in case of breach of integrity. I commit myself to providing the participant with a copy of this consent form.

Researcher's name, Surname, Date and Signature

*Note: Elements of the document concerning body odour collection have been removed ([...]), since they are not relevant to the present article.*

### Appendix C: Confidentiality agreement and terms of use for the Geneva Faces and Voices (GEFAV) database

Thank you for your interest in our GENEVA Faces And Voices (GEFAV) database. The GEFAV database is a series of faces and voices of males and females aged 18 to 35 years old. It entails pictures of neutral and smiling faces, video sequences of speaking and smiling faces, audio recordings of sentences and vowels. The GEFAV is being developed and distributed by the "Centre Interfacultaire en Sciences Affectives" at the University of Geneva. It is freely available exclusively for use in academic, non-profit research, in European countries and countries benefiting from a decision of the European Commission regarding adequacy of the protection of personal data.

*Note:* GEFAV stimuli have been controlled for their range of attractiveness, based on evaluations by a population with

the same characteristics as the stimuli. However, in no case should the GEFAV be considered as a standard database or as representing a universal range of human facial attractiveness. Especially, the distribution of attractiveness is partly a result of targeted recruitment of attractive individuals. Researchers using the GEFAV should thus keep in mind that the stimuli are appropriate for experimental studies, but may not be appropriate for correlational studies that assume an ecologically valid distribution of attractiveness.

**Collaborators who wish to use the GEFAV for their research have to fill out the agreement form below. By signing below the collaborator agrees to the following:**

- 1) That the GEFAV database remains the property of the University of Geneva and can be recalled at any time. Data from a specific participant can be recalled at any time on his/her request.
- 2) That the GEFAV database is for the collaborator's use only and is to be used only in the research project described below, except with the express permission of the "Centre Interfacultaire en Sciences Affectives" at the University of Geneva. Especially, it is required that the collaborator does not provide the database to a third party, especially to profit-making companies or organizations. Furthermore, the collaborator is required not to make the GEFAV available on the internet.
- 3) That the communications of any kind presenting data obtained with the help of the GEFAV systematically cite the following article: *Ferdenzi C., Delplanque S., Mehulblantar I., Da Paz Cabral K.M., Domingos Felicio, M., Sander D. (2014) The GENEVA Faces And Voices (GEFAV) database. Behavior Research Methods. doi:10.3758/s13428-014-0545-0*
- 4) That the "Centre Interfacultaire en Sciences Affectives" at the University of Geneva will be informed of any work performed using the database that has been submitted for publication to an academic journal.
- 5) **That this Confidentiality Agreement and Terms of Use is governed by the laws of Switzerland, excluding conflict of law rules, and that the competent jurisdiction is Geneva.**

**I agree to the above terms and conditions,**

---

**Date:** Collaborator's signature: Official signature of the institution and/or the Professor:

**Summary of the research:**

---

Name of the Collaborator (last name, first name):

Name of the Professor (last name, first name):

Academic Institution:

Department:

Address:

City:  
 Zip code:  
 State/Province:  
 Country:

Phone: Email (valid academic address only):

**Contacts:** Prof. David Sander (David.Sander@unige.ch)  
 or Dr Sylvain Delplanque (Sylvain.Delplanque@unige.ch),  
 Swiss Centre for Affective Sciences CISA, University of  
 Geneva, Case Postale 60, CH-1211 Geneva 20, Switzerland.

## References

- Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. *Emotion, 12*, 1161–1179. doi:10.1037/a0025827
- Barton, K. (2013). “MuMin: Multi-model interference” package: Model selection and model averaging based on information criteria (R Package Version 1.9.5). Retrieved from <http://cran.r-project.org/web/packages/MuMin/index.html>
- Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. (1997). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*, 711–720. doi:10.1109/34.598228
- Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*, 531–539. doi:10.3758/BRM.40.2.531
- Bestelmeyer, P. E. G., Latinus, M., Bruckert, L., Rouger, J., Crabbe, F., & Belin, P. (2012). Implicitly perceived vocal attractiveness modulates prefrontal cortex activity. *Cerebral Cortex, 22*, 1263–1270. doi:10.1093/cercor/bhr204
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.2.46). Retrieved from [www.praat.org](http://www.praat.org)
- Bogaert, A. F., Fawcett, C. C., & Jamieson, L. K. (2009). Attractiveness, body size, masculine sex roles and 2D:4D ratios in men. *Personality and Individual Differences, 47*, 273–278. doi:10.1016/j.paid.2009.03.011
- Braun, C., Gruendl, M., Marberger, C., & Scherber, C. (2001). Beautycheck: Ursachen und Folgen von Attraktivitaet (Technical report). Regensburg, Germany: Universität Regensburg. Available from [www.uni-regensburg.de/Fakultaeten/phil\\_Fak\\_II/Psychologie/Psy\\_II/beautycheck/english/bericht/bericht.htm](http://www.uni-regensburg.de/Fakultaeten/phil_Fak_II/Psychologie/Psy_II/beautycheck/english/bericht/bericht.htm)
- Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., & Belin, P. (2010). Vocal attractiveness increases by averaging. *Current Biology, 20*, 116–120. doi:10.1016/j.cub.2009.11.034
- Bryant, G. A., & Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters, 5*, 12–15. doi:10.1098/rsbl.2008.0507
- Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., & Weiss, B. (2005). *A database of German emotional speech*. Paper presented at Interspeech 2005, Lisbon, Portugal.
- Collins, S. A. (2000). Men’s voices and women’s choices. *Animal Behaviour, 60*, 773–780. doi:10.1006/anbe.2000.1523
- Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour, 65*, 997–1004. doi:10.1006/anbe.2003.2123
- DeBruine, L. M., Jones, B. C., Crawford, J. R., Welling, L. L. M., & Little, A. C. (2010). The health of a nation predicts their mate preferences: Cross-cultural variation in women’s preferences for masculinized male faces. *Proceedings of the Royal Society B, 277*, 2405–2410. doi:10.1098/rspb.2009.2184
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 285–290.
- Ekman, P., & Friesen, W. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour, 69*, 561–568. doi:10.1016/j.anbehav.2004.06.012
- Ferdenzi, C., Delplanque, S., Vorontsova-Wenger, O., Pool, E., Bianchi-Demicheli, F., & Sander, D. (2014). Perception of men’s beauty and attractiveness by women with low sexual desire. *Journal of Sexual Medicine, 11*. doi:10.1111/jsm.12795
- Ferdenzi, C., Lemaitre, J.-F., Leongomez, J. D., & Roberts, S. C. (2011). Digit ratio (2D:4D) predicts facial, but not voice or body odour, attractiveness in men. *Proceedings of the Royal Society B, 278*, 3551–3557. doi:10.1098/rspb.2011.0544
- Ferdenzi, C., Patel, S., Mehu-Blantar, I., Khidasheli, M., Sander, D., & Delplanque, S. (2013). Voice attractiveness: Influence of stimulus duration and type. *Behavior Research Methods, 45*, 405–413. doi:10.3758/s13428-012-0275-0
- Ferdenzi, C., Schaal, B., & Roberts, S. C. (2009). Human axillary odor: Are there side-related perceptual differences? *Chemical Senses, 34*, 565–571. doi:10.1093/chemse/bjp037
- Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *American Naturalist, 139*, 603–622. doi:10.1086/281919
- Gangestad, S. W., & Scheyd, G. J. (2005). The evolution of human physical attractiveness. *Annual Review of Anthropology, 34*, 523–548. doi:10.1146/annurev.anthro.33.070203.143733
- Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology, 108*, 233–242. doi:10.1037/0735-7036.108.3.233
- Gross, R. (2005). Face databases. In S. Li & A. Jain (Eds.), *Handbook of face recognition*. Berlin, Germany: Springer.
- Kelly, F., Drygajlo, A., & Harte, N. (2013). Speaker verification in score-ageing-quality classification space. *Computer Speech and Language, 27*, 1068–1084. doi:10.1016/j.csl.2012.12.005
- Kline, P. (1993). *Handbook of psychological testing*. London, UK: Routledge.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390–423. doi:10.1037/0033-2909.126.3.390
- Law Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., . . . Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society B, 273*, 135–140. doi:10.1098/rspb.2005.3296
- Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society B, 366*, 1638–1659. doi:10.1098/rstb.2010.0404
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *Karolinska Directed Emotional Faces—KDEF (CD ROM)*. Stockholm, Sweden: Karolinska Institutet, Department of Clinical Neuroscience, Psychology Section.
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36*, 630–633. doi:10.3758/BF03206543
- Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior, 25*, 229–241. doi:10.1016/j.evolhumbehav.2004.04.003
- Perrett, D. I. (2010). *In your face: The new science of human attraction*. Basingstoke, UK: Palgrave Macmillan.

- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior, 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., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887. doi:10.1038/29772
- Petta, P., Pelachaud, C., & Cowie, R. (2011). *Emotion-oriented systems: The Humaine handbook*. Berlin, Germany: Springer.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226. doi:10.1146/annurev.psych.57.102904.190208
- Ricciardelli, L. A., & Williams, R. J. (2012). Beauty over the centuries—Male. In T. F. Cash (Ed.), *Encyclopedia of body image and human appearance* (Vol. 1, pp. 50–57). San Diego, CA: Elsevier Academic Press.
- Roberts, S. C., & Little, A. C. (2008). Good genes, complementary genes and human mate preferences. *Genetica, 134*, 31–43. doi:10.1007/s10709-008-9254-x
- Roney, J. R., & Maestriepieri, D. (2004). Relative digit lengths predict men's behavior and attractiveness during social interactions with women. *Human Nature, 15*, 271–282. doi:10.1007/s12110-004-1009-5
- Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society B, 266*, 1913–1917. doi:10.1098/rspb.1999.0866
- Sentilles, R. M., & Callahan, K. (2012). Beauty over the centuries—Female. In T. F. Cash (Ed.), *Encyclopedia of body image and human appearance* (Vol. 1, pp. 43–49). San Diego, CA: Elsevier Academic Press.
- Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications, 21*, 42–50.
- Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion, 13*, 724–738. doi:10.1037/a0032335
- Vukovic, J., Jones, B. C., Feinberg, D. R., DeBruine, L. M., Smith, F. G., Welling, L. L. M., & Little, A. C. (2011). 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. *British Journal of Psychology, 102*, 37–48. doi:10.1348/000712610X498750