Received: 29 April 2022 | Accepted: 22 August 2022

DOI: 10.1002/ejsp.2905

EASP WILEY

RESEARCH ARTICLE

When faces and voices come together: Face width-to-height ratio and voice pitch contribute independently to social perception

Matteo Masi[1,2] | Simone Mattavelli[1] | Fabio Fasoli[2] | Marco Brambilla[1]

[1]University of Milano-Bicocca, Milano, Italy

[2]University of Surrey, Guildford, UK

Correspondence
Matteo Masi, University of Milano-Bicocca, Department of Psychology, Piazza dell' Ateneo Nuovo, 1, Milano 20126, Italy.
Email: m.masi8@campus.unimib.it

Abstract

Prior research has shown that the face's width-to-height ratio (fWHR) and the voice's pitch influence social perception. Yet, the relative contribution of either cue has been largely unexplored. We examined the simultaneous effects of fWHR and pitch on social evaluations. Experiment 1 ($N = 102$) tested how such cues shaped global impressions. Experiment 2 ($N = 121$) tested fWHR and pitch's effect on behavioural affiliative intentions, framing social interaction as a physical or an intellectual competition. Experiment 3 ($N = 57$) assessed whether variations in fWHR and pitch could influence trait attribution (i.e., physical formidability and intelligence). Individuals with large faces or low-pitched voices elicited negative impressions, positive behavioural intentions in a physical competition, and the attribution of stronger formidability but lower intelligence. Across the studies, cues exerted independent effects. The implications of these findings for research on cross-modal social perception are discussed.

KEYWORDS

behavioural intentions, cross-modal perception, face width-to-height ratio, formidability, intelligence, voice pitch

1 | INTRODUCTION

In everyday life, we base our judgements of others on their faces and voices. From facial and vocal features, people infer others' dispositions and intentions (McAller & Belin, 2018; Todorov et al., 2015). Face width-to-height ratio (fWHR) and the voice's pitch are two of the most-studied cues; indeed, they predict far-reaching consequences (Geniole et al., 2015; Pisanski & Bryant, 2019). For instance, large-faced individuals are perceived as untrustworthy and dominant (Mileva et al., 2014; Stirrat & Perret, 2010). In a similar vein, individuals consider low-pitched speakers as aggressive and untrustworthy (McAller et al., 2014; Zhang et al., 2021). Whereas the research substantially includes the unique contributions of both the face and the voice to shaping impressions and behavioural responses, it lacks investigation of their joint and relative impact on social perception. Here, we examined the concurrent effect of faces and voices, varying fWHR and pitch, respectively, across different dimensions and contexts. We focused on global impressions, behavioural intentions when choosing teammates, and dispositional inferences that may be relevant in specific contexts.

1.1 | Facial width-to-height ratio and social perception

A few seconds of exposure to a portrait are sufficient to infer many subject characteristics (Zebrowitz, 2017). The face is a powerful source that people spontaneously use to form an impression of others (Hehman et al., 2015a; Todorov et al., 2015). Among the many facial cues that people use to extract information and make inferences about the target individual, fWHR stands out as particularly relevant. The fWHR (i.e., the ratio of the bizygomatic width and the upper-face height) (Geniole et al., 2015) first appeared as an 'objective' measure of masculinity (Weston et al., 2007). Indeed, prior work has shown that fWHR correlates positively with testosterone level (Lefevre et al., 2013; but see Kordsmeyer et al., 2019) and predicts behavioural tendencies. A decade of research has suggested that fWHR serves as a criterion for mate selection and a means of increasing one's chances of winning inter-male competitions (i.e., Geniole et al., 2015; Puts et al., 2012). For instance, large-faced men are more aggressive hockey players than narrow-faced ones (Carré & McCormick, 2008). They exhibit great abilities in contact sports (Caton et al., 2022; Zilioli et al., 2015) or engage in antisocial behaviours to overcome other individuals (i.e., cheating) (Haselhuhn & Wong, 2012; Stirrat & Perret, 2010).

Although recent findings cast doubts on the ecological correlation between fWHR and such behaviours (i.e., Kosinski, 2017; Krenn and Meier, 2018; Wang et al., 2019), mounting evidence indicates that people nonetheless employ fWHR in social perception. For example, large-faced individuals are perceived as threatening (Geniole et al., 2015), aggressive, angry (Deska et al., 2018a), untrustworthy (Matsumoto and Hwang, 2021; Stirrat & Perret, 2010), and dominant (Mileva et al., 2014). Importantly, people use fWHR to make inferences about both the physical and mental characteristics of the target person. Hence, we see large-faced individuals as having greater muscularity and strength (i.e., formidability) (Zilioli et al., 2015). In contrast, these individuals are seen as less intelligent than the average person (Deska et al., 2018b). However, recent findings suggest that the impact of fWHR on dispositional inferences might be smaller and less reliable than previously thought, particularly when considering male faces (see Durkee and Ayers, 2021; Jaeger and Jones, 2020). Thus, the contribution of fWHR to person perception is still a matter of debate. Also, isolating the conditions under which fWHR influences dispositional inferences and attributions remains a prominent research goal. One way to establish the role of fWHR in social perception would be to test its effects in combination with other relevant and concurrent cues. Hence, we adopt this approach to studying the interplay between fWHR and vocal features.

1.2 | Perceiving others via voice pitch

People also judge others by their voices (Aronovitch, 1976). Listening to a few syllables is sufficient to form an impression about a speaker (McAller & Belin, 2018). Multiple vocal features distinguish one speaker from another and affect listeners' judgements (Pisanski and Bryant, 2019). Among those features, the voice's pitch (i.e., the perceptual outcome of the folds' vibration rate) (Fitch, 2000) relates to multiple outcomes. Just like fWHR, pitch has an association with testosterone level (Aung and Puts, 2020). Thus, low-pitched voices sound more masculine than high-pitched voices. Moreover, pitch negatively correlates with speakers' actual physical strength and performance (Aung et al., 2021; Puts & Aung, 2019). Despite the scarce and unpersuasive evidence showing that pitch is an accurate predictor of physical prowess (Armstrong et al., 2019; Feinberg et al., 2019; Pisanski et al., 2014; Schild & Zetter, 2021), listeners rely on vocal pitch to judge several speaker traits. For instance, low-pitched speakers are perceived as aggressive (Zhang et al., 2021), untrustworthy (McAller et al., 2014; but see Schild et al., 2020) and dominant (Puts et al., 2006). In addition, pitch predicts physical attribution—, that is, speakers with low-pitched voices are perceived as more formidable than those with high-pitched voices (Aung et al., 2021; Rendall et al., 2007). Hence, pitch implies several speaker characteristics that tend to overlap with those we infer from fWHR.

Importantly, most studies considered faces and voices in isolation. However, processing an individual's face and voice simultaneously occurs frequently in real-life encounters. Hence, studying fWHR and pitch in combination can provide further insights into the impact of each cue on social judgements.

1.3 | Simultaneous processing of fWHR and pitch

Most social perception studies have tested the impact of cues in isolation (Campanella and Belin, 2007). On many everyday occasions, perceivers process others' faces and voices independently, but multimodal perception in the social world is not unusual (McArthur and Baron, 1983). Indeed, our brain encodes facial and vocal cues through parallel pathways that connect throughout each elaboration stage (for a review, see Young et al., 2020). Thus, studying face-voice integration enables us to examine their relative and joint contribution to social perception.

Research has reflected a growing interest in the complementary impact of facial and vocal information (e.g., Hansen et al., 2017; Mileva et al., 2020; Zuckerman & Sinicopri, 2011). To date, studies show that faces and voices affect a range of traits independently, including dominance (Rezlescu et al., 2015), male attractiveness (Wells et al., 2013) and trustworthiness (Mileva et al., 2018). However, evidence also showed their interactive effects, such as on trustworthiness (Rezlescu et al., 2015) and female attractiveness (Wells et al., 2013). To the best of our knowledge, no study simultaneously presented faces and voices and tested their effect on global impressions. Masi et al. (2022) showed that faces can update initial global impressions of voices, and vice versa, but this finding applies to judgements resulting from the sequential presentation of each cue. Despite studies measuring the unique effects of fWHR and pitch on impressions (i.e., Han et al., 2017), no study systematically varied such features while presenting them simultaneously. Thus, whether fWHR and pitch processed simultaneously produce either independent or interactive effects on global impressions remains an unanswered question.

1.4 | From judgements to behavioural intentions

Impressions are just the first step in social interactions. Based on an ecological perspective of social perception, perceivers are attuned to others' features to prepare a proper behavioural reaction (McArthur and Baron, 1983). Hence, a better outline of effects of fWHR and voice pitch on social perception calls for investigating how they shape perceivers' behavioural intentions. Receptivity to cues' subtle meaning enables meeting the fundamental goal of self-preservation, that is, avoiding perceived danger and, in its absence, approaching (Chen and Bargh, 1999). In a typical situation, one keeps individuals with large faces and low-pitched voices at a distance, to decrease the chances of risking one's own life (see Lieberz et al., 2017). However, social judgements and behavioural intentions are often conditional on the circumstances in which one makes the judgement.

Real-life decisions often take place in contexts that may change the outcomes of interpersonal judgements. For example, Melnikoff and Bayley (2018) demonstrated that others' negative traits (i.e., immorality) can become desirable for teammate selection when participating in a game in which an immoral partner could constitute an advantage. The same reasoning could apply to behavioural intentions resulting from perceptions of fWHR and pitch. In interpersonal interactions, the conditions for the conventional rejection of individuals whose physical characteristics signal negative dispositions should be the context in which the interaction takes place. Since we commonly attribute superior physical abilities to large-faced individuals (Zilioli et al., 2015), they take precedence over narrow-faced individuals for tasks in which such a skill is appropriate (i.e., moving heavy weights) (Deska et al., 2018b; see also Hehman et al., 2015b). Similarly, when invited to choose a political representative, people prefer low-pitched rather than high-pitched speakers because greater leadership abilities are attributed to the former (Klofstad et al., 2012; Tigue et al., 2012). However, to our knowledge, no study assessed the contribution of such cues in eliciting perceivers' behavioural intentions when presented together. Therefore, fWHR and pitch should be tested in their simultaneous presentation to examine how people make use of such features to draw inferences and develop intentions as a function of the social contexts of decision-making.

1.5 | Present research

This research extended prior findings on face-voice integration for social perception by investigating the simultaneous occurrence of variations in fWHR and vocal pitch. We measured whether these two features elicit independent effects, interactive effects, or both in predicting impressions, behavioural intentions, and dispositional inferences. Experiment 1 tested the impact of fWHR and pitch on impression formation. We asked participants to express a global impression of a series of unknown individuals, each of which we associated with a specific face and voice presented simultaneously before the evaluation. The face could be either large or narrow, whereas the voice could be either low-pitched or high-pitched. Experiment 2 extended our investigation by looking at the interplay between fWHR and pitch in driving behavioural intentions across different contexts. We asked participants to indicate the extent to which they would choose individuals who varied in fWHR and pitch as fellow players in either a physical or intellectual competition. In Experiment 3, we examined the relationship between fWHR and pitch with perceived formidability and intelligence. Finally, we tested whether these attributions were key drivers of the contextual dependency of fWHR and pitch's effects on behavioural intentions.

2 | EXPERIMENT 1

Experiment 1 aimed to measure the contribution of both fWHR and pitch on global impression. In line with the extant literature, we expected both fWHR and pitch to influence impressions, with large (vs. narrow) faces and low- (vs. high-) pitched voices driving negative impressions. We also explored whether fWHR and vocal pitch led to independent effects or whether the two interacted in driving impressions.

Thus, we employed a 2(fWHR: large vs. narrow face) x 2(pitch: low vs. high) within-subjects design, asking participants to judge 12 individuals whose faces and voices we disclosed simultaneously.

2.1 | Method

2.1.1 | Participants

Sample size was based on prior work with a similar design. Rezlescu et al. (2015, Study 2) and Mileva et al. (2018, Studies 2–3) collected, respectively, 84, 64, and 80 participants. Thus, we collected a larger sample of 102 participants (77 females, $M_{age} = 26.59$, $SD_{age} = 9.77$).

We analysed the data by employing linear mixed models. We conducted a simulation-based sensitivity analysis (i.e., safeguard power analysis, Perugini et al., 2014). We used the "simr" R package to estimate the power of the critical effects using the 95% CI lower bounds in place of the observed value (Green and MacLeod, 2016). The estimated (unstandardized) effects of fWHR and pitch in Experiment 1 were .48 (lower bound was .38) and .12 (lower bound was .01), respectively. After replacing the observed effects with the relevant lower bound, the analysis ($n = 100$ simulations) yielded more than 99% of power for the fWHR effect, but 8% for the pitch effect (see OSF for the scripts of the simulations). The low power to detect this latter effect is examined in the general discussion.

2.1.2 | Procedure

We programmed the study in Qualtrics and recruited participants online through social networks. We asked respondents to participate in a study on impression formation, starting with putting on their headphones and adjusting the volume by playing an audio file (i.e., doorbell). The instructions told them that they would see pictures of individuals' faces at the centre of the monitor screen while voices played in the background. After each face-voice pair presentation (≈ 3000 ms), in place of the face, they saw the question, "What is your global impression of this person?" with a seven-point scale below (−3 = negative, 3 = positive).[1] Each participant received 12 face-voice pairs in random order. Despite no time limit on rating the target individuals, we reminded them to do it as fast as possible. At the end of the experiment, we asked participants to report their demographics (i.e., gender, age) and thanked and debriefed them.

2.1.3 | Materials

Faces

We extracted 12 high-quality photographs of male subjects, posing frontally and with a neutral expression, from the Chicago Face Database (CFD) (Ma et al., 2015) (see the OSF link for the chosen pictures' associated codes). To select the stimuli, we employed the fWHR measurement values reported with the database. We chose six large ($M = 2.04$, $SD = .07$) and six narrow faces ($M = 1.70$, $SD = .07$), $t(9.98) = 9.05$, $p < .001$, $d = 5.23$, 95% CI [2.68, 7.73].

Voices

Twelve Italian male speakers ($M_{age} = 23.10$, $SD_{age} = 1.44$) uttering the sentence, "I sat on the chair" ("Mi sono seduto sulla sedia", = ∼ 3000 ms) were recorded for this study. Sixty Italian independent raters evaluated the sentence as neutral in valence ($M_{age} = 27.52$, $SD_{age} = 8.21$, 45 females) with respect to the midpoint of a seven-point scale (−3 = negative, +3 = positive) ($M = 0.20$, $SD = 1.12$), $t(59) = 1.39$, $p = .17$, $d = .18$, 95% CI [−.08, .44]. Manipulation of speakers' voices created 12 high-pitched and 12 low-pitched voices by increasing or decreasing each speaker's average pitch ($M = 108.93$ Hz, $SD = 9.95$) by .50 Equivalent Rectangular Bandwidths (ERBs, about 20 Hz) using a Praat script (Feinberg, 2021). Research has proved such a manipulation to elicit a variation in the perception of several traits (Jones et al., 2010; O'Connor & Barclay, 2017). The resulting high-pitched voices ($M = 127.67$ Hz, $SD = 9.50$; $t(11) = 2.79$, $p = .02$, $d = 1.68$, 95% CI [.28, 3.03]) and low- ($M = 89.67$ Hz, $SD = 8.63$; $t(11) = -10.05$, $p < .001$, $d = 6.06$, 95% CI [3.28, 8.80]) were, respectively, above and below the average male mean pitch of 120 Hz (Gorris et al., 2020).

Face-voice combinations

We paired each face with a manipulated voice to constitute 72 different combinations (18 for each of the 2 × 2 table cells), half of which possessed the same face paired with the high- or low-pitched voice of each speaker.[2] Of these, 12 pairs (three large face-low voice, three narrow face-low voice, three large face-high voice, three narrow face-high voice) were randomly selected from the pool and appeared before each participant. These pairings were counterbalanced so that the same facial and (manipulated) vocal identities never appeared to participants twice in the same session.

TABLE 1 Experiment 1. Means and standard deviations of global impressions as a function of the combinations of fWHR (narrow vs. large face) and pitch (high vs. low voice)

	High-pitched voice		Low-pitched voice	
	M	*SD*	*M*	*SD*
Narrow face	.72	.48	.49	1.53
Large face	−.25	1.45	−.50	1.55

Scores ranging from −3 (negative to 3 (positive).

2.2 | Statistical analysis

In all the experiments, we analysed results with linear mixed models in R, using the 'lmerTest::lmer()' function (Kuznetsova et al., 2017), and estimated marginal means with 'emmeans::emmeans()' (Lenth, 2021). We tested several model solutions (e.g., random intercepts and slopes, simplified random correlations) (see Barr et al., 2013). We chose to report the model without singularity/convergence issues. All experiments corresponded to the model with the participant and the face–voice combination identifiers inputted as random intercepts but without random slopes. See Supplementary Materials for details on random structure and models fit.

Significant interactions were inspected with simple effects analysis. Since no factor presented more than two levels, a single t-test was performed within each level of the other factors (or factorial combinations). Indeed, no multiplicity issue was created, eliminating the necessity for the p-value adjustment.

2.3 | Results and discussion

fWHR, pitch and their interaction were effect-coded and introduced as fixed factors. Means and standard deviations are reported in Table 1. We found a significant effect of fWHR showing that narrow-faced individuals were evaluated more positively than large-faced individuals, $B = .49$, $SE = .06$, $t(65.20) = -9.15$, $p < .001$. Similarly, a significant effect of pitch indicated that high-pitched speakers were judged more positively than low-pitched speakers, $B = .12$, $SE = .04$, $t(65.20) = -2.22$, $p = .03$. The interaction between fWHR and pitch was not significant, $B = -.01$, $SE = .03$, $t(65.20) = -.18$, $p = .85$. The main effect of fWHR was significantly bigger than the effect of pitch, $B = .37$, $SE = .07$, $Z = -4.98$, $p < .001$.

Thus, both fWHR and vocal pitch influenced impression formation independently from each other, such that large (vs. narrow) faces and low- (vs. high-) pitched voices lead to negative impressions of the targets. Notably, there was no interaction between the two cues.

3 | EXPERIMENT 2

We designed Experiment 2 to extend Experiment 1 by examining behavioural intentions that fWHR and pitch elicited and considering

[1] Participants were also shown two alternated and counterbalanced blocks in which they evaluated each target on masculinity (not masculine vs. masculine) and threat (threatening vs. non-threatening) on seven-point scales. Results appear in Supplementary Materials.

[2] We decided to manipulate pitch but not fWHR, for two reasons. First, a pilot study ($N = 18$) showed that a manipulation of face width (see Hehman et al., 2015b) was not sufficient to elicit an effect of fWHR (at least in our sample). Second, our speakers did not show audible differences in terms of the average pitch when not manipulated, requiring a digital modification to obtain high- and low-pitched voices.

the role of the decision context. Perceivers used others' fWHR and pitch as indicators of physical strength (Puts et al., 2012). However, the ultimate impact of such cues on the perceiver's judgements might depend on the specific context of such decision-making (Hehman et al., 2019). Critically, large faces and low-pitched voices could become desirable features in a situation in which physical strength becomes functional for achieving a specific goal. For instance, people involved in a conflict could approach rather than avoid individuals showing such features (Van Vugt, 2009). Indeed, affiliating with someone imagined as more muscular or perceived as more intimidating—both attributes that fWHR and pitch cue (Deska et al., 2018a; Hehman et al., 2013; Puts et al., 2016)—is likely to increase chances of victory over adversaries. As a case in point, Hehman et al. (2015b) gave participants the goal of choosing teammates to win a competition. They also varied (i) targets' fWHR (i.e., large or narrow) and (ii) context (i.e., physical vs. intellectual competition). They found that participants selected large-faced persons to form a team to win a physical competition, but the reverse for an intellectual competition.

We adopted a similar procedure to test whether both fWHR and pitch simultaneously appearing predicted such behavioural intentions (i.e., affiliative intentions) according to the context (i.e., competition type). The experiment consisted of a 2 (fWHR: narrow vs. large) × 2(pitch: high vs. low) × 2 (competition: physical vs. intellectual) within-subjects design. We envisioned participants choosing large- (vs. narrow-) faced individuals more often as teammates for winning a physical competition (i.e., boxing) rather than an intellectual competition (i.e., chess, replicating Hehman et al., 2015b). Furthermore, and in line with the idea that fWHR and pitch serve similar purposes, we hypothesised the type of competition to qualify the impact of vocal pitch on behavioural intentions. That is, vocally low-pitched (vs. high-pitched) individuals would be chosen more often as teammates for a physical competition than an intellectual competition. As in Experiment 1, we explored the presence of independent and interaction effects.

3.1 | Method

3.1.1 | Participants

We collected a sample of 121 participants (93 Females, M_{age}= 24.68, SD_{age} = 9.81) on social networks with a direct link to Qualtrics. We employed a safeguard power analysis to provide evidence on this experiment's power. Unlike Experiment 1, here we focused on the two interaction effects (i.e., fWHR*Competition and Pitch*Competition). The estimated (unstandardized) fWHR*Competition effect was −.53 (lower bound was −.47), whereas the Voice Pitch*Competition effect was −.18 (lower bound was −.12). After replacing the observed effects with the respective lower bound the analysis (n = 100 simulations) always yielded more than 99% of power.

3.1.2 | Procedure

The experiment was programmed on Qualtrics. We asked participants to choose their partners for either a boxing or a chess team (physical vs. intellectual competition). The two conditions appeared in alternated blocks. Thus, the instructions were:

> You will create your team for a boxing [chess] competition. Your goal is to create the best team to defeat your adversaries in a tournament. You will be presented with one candidate per time. For each, you will be presented with his face and his voice only. Right after each presentation, you will be asked to express how much you want each person to become part of your team for winning the competition.

In each block (boxing vs. chess), we presented participants with 12 unique face-voice pairs in random order (≈ 3000 ms). After each pair, a seven-point scale at the centre of the screen asked for a rating of participants' behavioural (i.e., affiliative) intentions: "To what extent do you want this person to become part of your team for winning the boxing [chess] competition?" (1 = not at all; 7 = definitely). Finally, we asked them to report their age and gender and debriefed them.

3.1.3 | Materials

Faces

We extracted 24 male faces from the CFD (Ma et al., 2015). In terms of fWHR, the 12 large faces (M = 2.00, SD = .04) were larger than the 12 narrow faces (M = 1.70, SD = .04), t(21.95) = 16.53, p < .001, d = 7.06, 95% CI [4.81, 9.28].

Voices

We sampled 12 new Italian male speakers (M_{age}= 25.46, SD_{age} = 4.21) uttering the Experiment 1 sentence. Voices underwent the same transformation to create their low- and high-pitched variations. Then, we added them to the set of 24 voices we created in Experiment 1, forming a set of 48 voices—*low-pitched voices* (M = 92.92 Hz, SD = 15.17); *high-pitched voices* (M = 131.57 Hz, SD = 17.74); t(44.92) = 8.11, p < .001, d = 2.41, 95% CI [1.64, 3.18].

Face-voice combinations

We paired faces and voices to constitute 192 unique combinations (48 for each of the 2 × 2 table cells), counterbalancing them among participants such that the same facial and (manipulated) vocal identities never appeared twice in the same session.

TABLE 2 Experiment 2. Means and standard deviations of behavioural (i.e., affiliative) intentions as a function of the combinations of the competition type (physical vs. intellectual), fWHR (narrow vs. large face), and pitch (high vs. low voice)

	Physical competition				*Intellectual competition*			
	High-pitched voice		Low-pitched voice		High-pitched voice		Low-pitched voice	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Narrow face	3.01	1.73	3.73	1.73	5.00	1.62	4.89	1.61
Large face	4.38	1.75	4.92	1.69	4.16	1.68	4.08	1.79

Scores ranging from 1 (not at all) to 7 (definitely).

3.2 | Results and discussion

Behavioural intentions towards targets were analysed as dependent variable. fWHR, pitch, and competition were effect-coded and employed as fixed factors. We reported means and standard deviations in Table 2.

We found the main effects of fWHR, B = -.11, SE = .04, $t(183) = -2.62$, $p = .009$, and pitch, B = -.13, SE = .04, $t(183) = -3.03$, $p = .003$. Regardless of the competition type, individuals with large faces and low-pitched voices elicited more positive behavioural intentions than those with narrow faces and high-pitched voices. Moreover, the main effect of competition type was significant, B = -.26, SE = .03, $t(2629) = -8.64$, $p < .001$. On average, individuals were chosen to a greater extent as partners for a chess than for a boxing competition. The interaction between fWHR and competition was significant, $B = -.52$, $SE = .03$, $t(2629) = -17.54$, $p < .001$. Direct contrasts showed that targets with large faces elicited more positive behavioural intentions in the boxing competition than those with narrow faces, $B = -1.28$, $SE = .10$, $t(397) = -12.20$, $p < .001$, but the reverse was found for the chess competition, $B = .82$, $SE = .10$, $t(397) = 7.89$, $p < .001$. Furthermore, the interaction between pitch and competition was significant, $B = -.18$, $SE = .03$, $t(2629) = -6.00$, $p < .001$, showing that targets with low- (vs. high-) pitched voices elicited more positive behavioural intentions for the boxing competition, B = -.62, SE = .10, $t(397) = -5.93$, $p < .001$, but no difference occurred in the chess competition, $B = .10$, $SE = .10$, $t(397) = .95$, $p = .34$. We also compared the two interactions to explore which factor between fWHR and pitch showed the larger difference between the chess and the boxing competition's intentions. fWHR contributed more than pitch in determining behavioural intentions towards targets, B = -.34, SE = .04, $Z = -8.15$, $p < .001$. No other effects were significant, $p > .28$.

Results demonstrated that, as in Experiment 1, fWHR and pitch affected inferences independently. Individuals with large (vs. narrow) faces elicited more positive behavioural intentions, that is being chosen as fellows, for a prospective physical competition. The reverse occurred for the intellectual competition. Moreover, the target's voice predicted positive behavioural intentions only within the physical competition, such that targets with a low- (vs. high-) pitched voice were more likely to be chosen.

4 | EXPERIMENT 3

[Experiment 2 clarified that the (simultaneous) impact of fWHR and vocal pitch on behavioural intentions to affiliate with a partner in a competition depended on the nature of the competition. Yet, these findings remain silent with respect to the inferences about target dispositional characteristics that both fWHR and pitch elicit and that influence perceivers' ultimate decisions. To address this question, we explored the impact of fWHR and pitch on perceived formidability (i.e., physical strength, muscularity, pain resistance) and intelligence. The experiment consisted of a 2 (fWHR: narrow vs. large) × 2 (pitch: high vs. low) × 2(individual disposition: formidability vs. intelligence) within-subjects design. Our prior findings suggested that the two cues would be independent predictors. Specifically, we expected fWHR to predict both attributions: that is, large (vs. narrow) faces should increase the attribution of formidability (Zilioli et al., 2015) but decrease the attribution of intelligence (Deska et al., 2018b). Similarly, low- (vs. high-) pitched voices should increase the formidability attribution (Aung et al., 2021), while their role in inferring intelligence should be only marginal (see Schroeder and Epley, 2016). Hence, we explored whether attribution of formidability and intelligence as a function of fWHR and pitch were associated with behavioural intentions observed in Experiment 2.

4.1 | Method

4.1.1 | Participants

We recruited 57 participants (46 females, M_{age}= 22.96, SD_{age} = 8.65) on social networks with a direct link to Qualtrics. A safeguard power analysis provided evidence on this experiment's power. We focused on the two interaction effects (i.e., fWHR*Disposition and Pitch*Disposition). The estimated (unstandardized) effect fWHR*Disposition interaction was −.34 (lower bound was −.30), and the Pitch*Disposition interaction was −.09 (lower bound was −.06). After replacing the observed effects with the respective lower bound the analysis (n = 100 simulations) always yielded more than 88% of power.

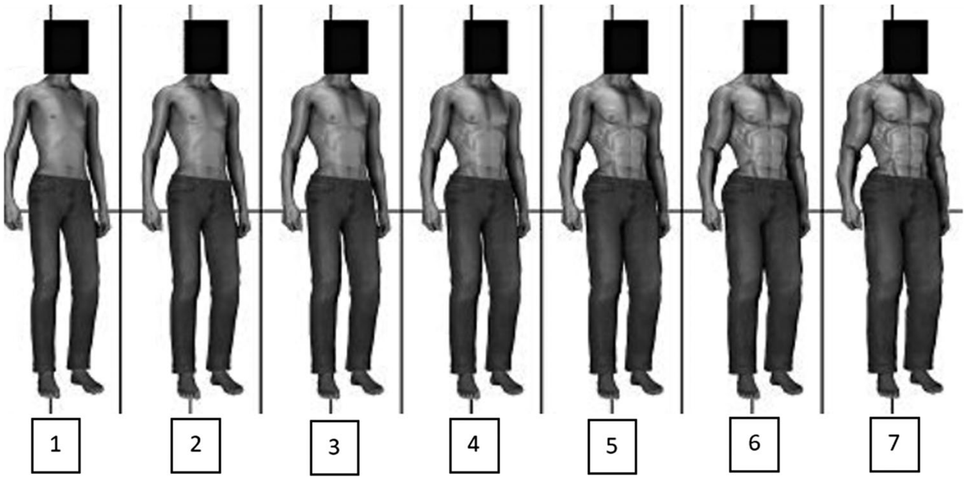

FIGURE 1 Experiment 3. Measure of muscularity consisting in a 'body matching' task. Participants chose the body of each individual target, presented beforehand with his face and voice, among pictures depicting male bodies varying in muscularity from 1 (low muscularity) to 7 (high muscularity). Adapted with permission from Frederick and Peplau (2007).

4.1.2 | Procedure

We presented participants with 24 unique targets extracted from the set of 192, equally distributed in two alternated blocks randomised in their order, one about perceived formidability and one about perceived intelligence. In each block, participants perceived each target through face-voice simultaneous presentation (≈ 3000 ms). After each, we asked participants to rate each target on either formidability or intelligence. At the end, they reported their demographics (i.e., age and gender) and we debriefed them.

4.1.3 | Materials

Face-voice combinations

We employed the same set of 192 face-voice combinations from Experiment 2.

Formidability

Two items assessed strength (i.e., "To what extent do you think this person can be strong?") and pain resistance (i.e., "To what extent do you think this person can be resistant to pain?"), respectively. Participants answered using a scale from 1 (*not at all*) to 7 (*very much*). For a third item—a body-matching question—we asked participants to choose the target's body from among seven pictures of increasingly muscular male bodies (Frederick and Peplau, 2007; Wilson et al., 2017) (see Figure 1). Reliability was high (Cronbach's $\alpha = .92$).

Intelligence

We administered three items to assess perceived intelligence (i.e., "To what extent do you think this person can be intelligent?"; "To what extent do you think this person can solve a complex mathematical problem?"; "What can be the intelligence quotient of this person?"). Participants provided their responses on a seven-point scale (i.e., 1 = not at all; 7 = very much, for the first two questions; for the third question, 1 = very low; 7 = very high). Reliability was high (Cronbach's $\alpha = .92$).

4.2 | Results and discussion

We treated fWHR, pitch, and individual disposition as fixed factors (effect-coded). We reported means and standard deviations in Table 3.

We found a significant main effect of individual disposition, such that attributions of perceived intelligence were stronger than those on formidability, $B = -.31$, $SE = .02$, $t(3996) = -16.21$, $p < .001$. The main effect of fWHR was significant, $B = -.15$, $SE = .03$, $t(182) = -4.44$, $p < .001$. Overall, large faces lead to stronger attributions than narrow faces. The interaction between fWHR and individual disposition was significant, $B = -.34$, $SE = .02$, $t(3996) = -17.67$, $p < .001$. As expected, large (vs. narrow) faces lead to stronger attributions of formidability, $B = -.99$, $SE = .08$, $t(314) = -12.67$, $p < .001$, and the pattern was reserved for intelligence, $B = .38$, $SE = .08$, $t(314) = 4.98$, $p < .001$. Also, the interaction between pitch and individual disposition was significant, $B = -.10$, $SE = .02$, $t(3996) = -5.00$, $p < .001$. Low- (vs. high-) pitched voices led to stronger attributions of formidability, $B = -.23$, $SE = .08$, $t(314) = -3.01$, $p = .003$, while a marginally significant effect was found for perceived intelligence, $B = .15$, $SE = .08$, $t(314) = 1.98$, $p = .05$, indicating that high- (vs. low-) pitched voices tended to lead to attributions of higher intelligence. We also compared the two interaction effects to explore which cue was more likely to draw both attributions: fWHR appeared to produce more extreme judgements than pitch, $B = -.25$, $SE = .03$, $Z = -8.95$, $p < .001$. No other effect or interaction was significant, $p > .08$.

TABLE 3 Experiment 3. Means and standard deviations of the averaged ratings of inferred formidability and intelligence as a function of fWHR (narrow vs. large face) and pitch (high vs. low voice)

	Formidability				*Intelligence*			
	High-pitched voice		Low-pitched voice		High-pitched voice		Low-pitched voice	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Narrow face	2.99	1.36	3.09	1.27	4.49	1.46	4.25	1.48
Large face	3.82	1.46	4.23	1.36	4.01	1.26	3.93	1.22

Scores ranging from 1 (low) to 7 (high).

Findings showed that participants ascribed higher formidability to targets with large faces and low-pitched voices. These same targets were also perceived as lacking intelligence. Consistent with our hypothesis, cues created independent effects in determining inferences of both formidability and intelligence.

5 | EXPLORATORY MEDIATION ANALYSIS

We explored whether the Experiment 3 ratings about perceived formidability and intelligence could explain why fWHR and pitch triggered specific behavioural intentions in Experiment 2, which diverged according to the type of competition.

5.1 | Data preparation

Experiment 3 data on the impact of fWHR and pitch on perceived formidability and intelligence ratings were averaged on the levels of the 192 face-voice combinations. Thus, each face-voice combination received two average scores, one on formidability and one on intelligence. Then, we prepared a new dataset merging such average scores with Experiment 2 data on behavioural intentions. In other words, we used the average scores as normative values of each face-voice combination collected on another sample with respect to Experiment 2 (for a comparable analytic procedure, see Deska et al., 2018b).

To explore whether formidability and/or intelligence mediate the fWHR and pitch effects on behavioural intentions, we built four mediation models (see Figure 2 for graphical representations). We separately examined participants' behavioural intentions concerning the two types of competition (physical vs. intellectual) and treated fWHR and pitch as independent predictors in all models. We also treated perceived formidability and intelligence as independent mediators. When not incurring singularity warnings, we employed the participant and the face-voice combination identifiers as random intercepts (as in the previous experiments); otherwise, we dropped the latter. We inspected indirect effects with 2000 bootstrap replicates (using more was impractical, due to computation limitations). We do not reiterate results that overlap with those previously reported.

5.2 | Results and Discussion

5.2.1 | Physical competition

Perceived formidability predicted participants' behavioural intentions, $B = .66$, $SE = .08$, $t(186) = 8.03$, $p < .001$, whereas intelligence did not, $B = -.12$, $SE = .09$, $t(189) = -1.23$, $p = .22$. Importantly, formidability mediated the effect of both fWHR (−.29, 95% CI [−.34, −.25]) and pitch (−.09, 95% CI [−.12, −.07]), such that individuals with large faces or low-pitched voices were perceived as more formidable and target of more positive behavioural intentions.

5.2.2 | Intellectual competition

Perceived formidability, $B = -.27$, $SE = .08$, $t(194) = 3.44$, $p < .001$, and intelligence $B = .24$, $SE = .08$, $t(189) = 3.04$, $p = .003$, predicted participants' behavioural intentions. Perceived intelligence mediated the fWHR effect (.05, 95% CI [.02, .07]), such that individuals with narrow faces were perceived as more intelligent and associated with more positive behavioural intentions. Similarly, there was an indirect effect of intelligence for the pitch effect on behavioural intentions but approaching the inclusion of the null value in the confidence intervals (.02, 95% CI [.007, .03]). Perceived formidability mediated the fWHR effect (.12, 95% CI [.07, .17], such that targets with large faces were perceived as more formidable and therefore elicited negative behavioural intentions. Also, even though the direct effect of pitch on behavioural intentions was not significant, formidability mediated its effect (.04, 95% CI [.02, .05]), such that targets with low voices were perceived as more formidable and thus elicited negative behavioural intentions.

These analyses showed that fWHR (from narrow to large) and pitch (from high to low) predicted positive inferences on formidability which, in turn, explained why targets with larger faces and lower voices were more likely to elicit positive behavioural intentions for a competition based on physical abilities. At the same time, fWHR negatively predicted intelligence, and this effect explained negative behavioural intentions in a logic-based game. Moreover, formidability explained the relationship between fWHR or pitch on behavioural intentions in an intellectual competition. While we consider these results informative, we need to be careful with their interpretation as these are exploratory

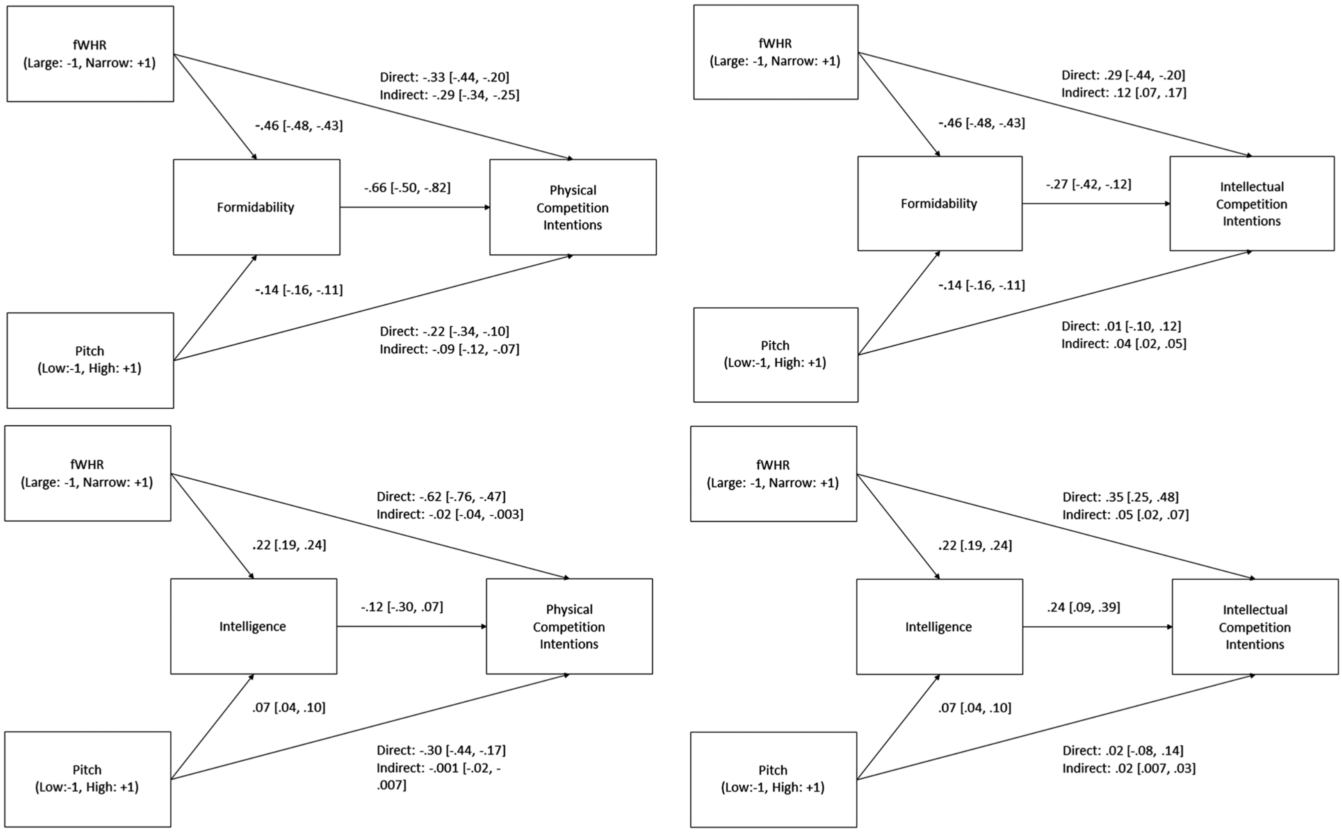

FIGURE 2 Exploratory Mediation Analysis. Graphical representations of the four mediation models. fWHR and pitch were treated as predictors, formidability and intelligence as mediators, and behavioural intentions in the physical and the intellectual competitions as dependent variables. Unstandardised effects and 95% confidence intervals are reported.

analyses performed on data occurring from two different experiments merged into one.

6 | GENERAL DISCUSSION

Research concerning fWHR and pitch has shown that the two cues share social functions (e.g., Puts et al., 2012), establishing them as predictors of social outcomes (e.g., Aung and Puts, 2020; Geniole et al., 2015). Yet, researchers mostly examined the impact of these two cues with faces and voices that perceivers processed in isolation. Filling this gap, we tested the impact of faces and voices varying in fWHR and pitch, respectively, in conditions of simultaneous processing. Hence, we assessed their relative contribution to predicting global impressions, behavioural intentions and dispositional characteristics. In Experiment 1, we found that fWHR and pitch were independent contributors to global impressions (i.e., they did not interact with one another). Participants evaluated individuals with large faces or low-pitched voices more negatively than those with narrow faces and high-pitched voices, respectively. In Experiment 2, we moved a step further and examined behavioural intentions (i.e., willingness to choose a target as a teammate to win a competition), by also varying the nature of the context for expressing such intentions (i.e., physical vs. intellectual competition). In a physical competition, fWHR and pitch independently predicted participants' intentions for choosing a teammate. Individuals with large faces or low-pitched voices were more likely to be choices for teammates than those with a narrow face or a high-pitched voice. The reverse pattern appeared in the context of an intellectual competition, but only fWHR affected participants' behavioural intentions. Building on this, Experiment 3 investigated the nature of participant inferences made on the target's dispositional characteristics, due to both fWHR and pitch variations that could partially explain Experiment 2 outcomes. We found that participants perceived individuals with large faces or low-pitched voices as more formidable but less intelligent than those with a narrow face or a high-pitched voice, respectively.

Our findings offer novel insights for research on cross-modal social perception. Prior work has shown that when vocal and facial cues appear together, they can elicit either independent or interaction effects (e.g., Mileva et al., 2018; Rezlescu et al., 2015; Wells et al., 2013). The present investigations looked differently at the combination of two specific features, often driving the social perception of both the face and the voice (i.e., fWHR and pitch). To our knowledge, this is the first study measuring global impressions of fWHR and pitch in simultaneous presentations. We found that both cues affect global impressions. Notably, it happened always with independent effects of either cue. This finding aligns with previous cross-modal perception studies that found independent effects of other facial and vocal features on dominance (Mileva et al., 2018; Rezlescu et al., 2015), trustworthiness (Mileva et al., 2018, but see Rezlescu et al., 2015) and attractiveness attributions to male targets (Wells et al., 2013). Still, the non-significant interaction between facial and vocal features observed in the current studies should be interpreted cautiously and limited to the specific manipulations employed in our research. Relatedly, the safeguard power analysis from Experiment 1 indicated that our sample was adequately powered to detect an effect of fWHR, but not of pitch, on global impressions. Indeed, the more substantial impact of fWHR does not necessarily call for a prominent role of this facial feature over voice pitch in altering impressions. As we could not employ a standardized manipulation of either cue, any direct comparison of the relative impact of fWHR and voice pitch would be premature at this stage. Yet, alternative ways to manipulate vocal pitch should establish (i) its relative contribution to influencing impressions of others and (ii) its interplay with fWHR. Moreover, future studies should substantiate whether fWHR and pitch also do not interact in making other trait-based judgements. Likewise, our findings provide a starting point for examining other social inferences. For instance, one could examine whether and how the two cues interact in inducing obedience/compliance in others (i.e., social influence), or how they affect trust in others' statements (i.e., truth attribution).

Our research was not limited to the investigation of cross-modal impressions, but also measured behavioural (i.e., affiliative) intentions as a function of salient contexts of decision (i.e., physical vs. logical/intellectual), which were supposed to activate such cues' social functions. We showed that fWHR determined intentions (as in Hehman et al., 2015b), as pitch did in the case of a physical competition, and both cues informed perceivers about target formidability. Moreover, our research adds knowledge on the relative role of fWHR and pitch on the target's perceived intelligence, a dispositional characteristic that has received little attention in prior studies. We not only replicated previous findings showing the role of fWHR in affecting behavioural intentions in intelligence-related contexts (i.e., Deska et al., 2018b; Hehman et al., 2015b). We also showed that pitch had a marginal influence on intelligence attribution (see also Schroeder and Epley, 2016). Once again, on all occasions, the simultaneous presentation of fWHR and pitch did not yield an interaction effect but two parallel and independent effects. Our results also established that context matters in determining the relative impact of fWHR and vocal pitch on behavioural intentions. Thus, researchers could investigate alternative instances of the functional roles of fWHR and pitch overlapping, as in physical competition, and when they do not, as in intellectual competition.

At present, we can advance a potential explanation concerning the independence of fWHR and pitch effects. fWHR might signal immutable dispositions (i.e., one can barely regulate one's own facial shape and musculature if not making emotional expressions or tilting the head) (Hehman et al., 2013; Witkower and Tracy, 2019). Conversely, one can effortlessly modulate one's own pitch and subject it to greater variability (Pisanski et al., 2016). Therefore, this latter cue might convey information that varies over time (i.e., one can lower the voice to sound more or less dominant) (Puts et al., 2006; see also Armstrong et al., 2019; Cheng et al., 2016; Zhang et al., 2021). Indeed, participants might have thought that speakers could voluntarily adjust pitch, to signal, for instance, potential formidability, while the speaker truly is not as formidable as the voice indicates. Thus, whereas fWHR signals the general disposition of a person, pitch may signal a transient status that the target can deliberately alter. The information that fWHR and pitch convey might differ, which could explain their independent effect on impressions. Similar reasoning has been proposed elsewhere (see Wells et al., 2013). To date, we do not know whether people hold lay beliefs concerning fWHR and pitch—fWHR signalling a disposition, pitch a temporary state—that may explain why the two were independent contributors. Considering the part that physiognomic beliefs play in determining how people rely on facial features when judging others (Jaeger et al., 2018; Suzuki et al., 2019), testing whether beliefs concerning fWHR and pitch variations affect appraisals and combinations would be intriguing.

Before concluding, we underline at least two methodological issues that could have affected our results and for which upcoming work should account. First, to obtain a more precise measurement of each cue's relative contribution, one should compare face-only and voice-only ratings to ratings when presented simultaneously. Indeed, it would constitute a more powerful test of their independence. Second, a critical reader might ask whether we ascribe the current findings to participants' inclination to believe that specific faces and voices belonged to the same person. In fact, to expand the range of targets and, consequently, the number of trials, we randomly combined faces and voices taken from different persons. Even though instructions presented face-voice pairs as belonging to the same person, mismatched facial and vocal identities might have affected participants' evaluations. Considering people's ability to match others' unfamiliar faces and voices beyond chance (i.e., Kamachi et al., 2003; Mavica and Barenholtz, 2013), the information coming from faces and voices may have not been easy to combine because participants found it hard to believe that the two cues belonged to the same identity. Eventually, this could have facilitated the occurrence of two independent effects rather than a potentially multiplicative/interactive one. Relevant prior work on person distinctiveness judgements (Tatz et al., 2020) found evidence of a considerable interaction between face and voice cues affecting person perception. Thus, future work should replicate our findings and investigate the conditions that may enhance the independence of fWHR and pitch effects, also considering such methodological issues.

This work also has other limitations. First, we inspected the role of fWHR and vocal pitch considering only male targets because the literature identified fWHR and pitch as sexually dimorphic characteristics peculiar to inter-male contests (Puts et al., 2012; but see Kramer, 2017, for a different position concerning fWHR). However, extending the investigation to female targets could be highly informative, as well as comparing ratings of male and female participants (see Wells et al., 2013). Second, we manipulated speakers' pitch to increase the difference between low-pitched and high-pitched voices, but we did not manipulate fWHR. The results of a pilot study suggesting that an attempted manipulation of facial width was not effective (inspired by Hehman et al., 2015b) motivated our decision. This implies that additional facial features (e.g., adiposity, skin quality) characterised unmanipulated faces that tend to naturally covary with fWHR and might ultimately have contributed to the effect of fWHR. Instead, manipulating voices implies that such a manipulation left unaffected those features naturally covarying with vocal pitch (e.g., timbre). Under such conditions, perceptions of face and voice stimuli might have differed, in terms of being natural (unmanipulated) or artificial, respectively. This aspect is especially relevant if one considers that subjects perceiving stimuli as less common in their daily environment (e.g., artificial faces) trust them less (Balas and Pacella, 2017). Future investigations should use more sophisticated manipulation methodologies (i.e., DeBruine, 2018; Stirrat & Perret, 2010) to assess the relative contribution of manipulated faces and voices to social perception.

7 | CONCLUSION

This work provided a first empirical investigation of the functions and the contribution of two well-established facial and vocal features, fWHR and pitch, by testing them simultaneously for social perception. Such cues showed independent contributors for several judgements, ranging from global impressions to behavioural intentions and including inferences on target characteristics (i.e., formidability and intelligence). Our effort provides new evidence to the growing field of cross-modal social perception. More importantly, this work could become an essential source for what may turn out to be a variety of investigations into how face/voice cues combine to produce effects on person perception, behavioural intentions, and other relevant social outcomes. Thus, researchers should expand such examination to further facial and vocal features and test the conditions favouring face-voice combination or independence.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All data, materials, and analyses code are available at https://osf.io/sm2j7/.

ETHICS STATEMENT

The research was conducted in accordance with the Declaration of Helsinki. Approval was obtained by the local university ethics committee. Participants provided their consent before taking part in the study.

TRANSPARENCY STATEMENT

This manuscript is an honest, accurate, and transparent account of the studies being reported; no important aspects of the studies have been omitted; any discrepancies from the studies as planned have been explained.

ORCID

Matteo Masi https://orcid.org/0000-0001-6436-1663
Simone Mattavelli https://orcid.org/0000-0002-8934-8016
Fabio Fasoli https://orcid.org/0000-0002-8043-1630
Marco Brambilla https://orcid.org/0000-0002-4774-3309

REFERENCES

Armstrong, M. M., Lee, A. J., & Feinberg, D. R. (2019). A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour*, *147*, 43–51. https://doi.org/10.1016/j.anbehav.2018.11.005

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, *99*(2), 207–220. https://doi.org/10.1080/00224545.1976.9924774

Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, *33*, 154–161. https://doi.org/10.1016/j.copsyc.2019.07.028

Aung, T., Rosenfield, K. A., & Puts, D. (2021). Male voice pitch mediates the relationship between objective and perceived formidability. *Evolution and Human Behavior*, *42*(2), 121–129. https://doi.org/10.1016/j.evolhumbehav.2020.08.007

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Balas, B., & Pacella, J. (2017). Trustworthiness perception is disrupted in artificial faces. *Computers in Human Behavior*, *77*, 240–248. https://doi.org/10.1016/j.chb.2017.08.045

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–543. https://doi.org/10.1016/j.tics.2007.10.001

Campbell, J. I., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*, *44*(4), 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651–2656. https://doi.org/10.1098/rspb.2008.0873

Caton, N. R., Hannan, J., & Dixson, B. J. W. (2022). Facial width-to-height ratio predicts fighting success: A direct replication and extension of Zilioli et al. (2014). *Aggressive Behavior*, *48*(5), 449–465. Portico. https://doi.org/10.1002/ab.22027

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*(2), 215–224. https://doi.org/10.1177/2F0146167299025002007

Cheng, J. T., Tracy, J. L., Ho, S., & Henrich, J. (2016). Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans. *Journal of Experimental Psychology: General*, *145*(5), 536–547. https://doi.org/10.1037/xge0000166

Clarkson, T. R., Sidari, M. J., Sains, R., Alexander, M., Harrison, M., Mefodeva, V., & Dixson, B. J. (2020). A multivariate analysis of women's mating strategies and sexual selection on men's facial morphology. *Royal Society Open Science*, *7*(1), 191209. http://doi.org/10.1098/rsos.191209

DeBruine, L. (2018). debruine/webmorph: Beta release 2 (v0.0.0.9001). https://doi.org/10.5281/zenodo.1162670

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018a). The face of fear and anger: Facial width-to-height ratio biases recognition of angry and fearful expressions. *Emotion*, *18*(3), 453–464. https://doi.org/10.1037/emo0000328

Deska, J. C., Lloyd, E. P., & Hugenberg, K. (2018b). Facing humanness: Facial width-to-height ratio predicts ascriptions of humanity. *Journal of Personality and Social Psychology*, *114*(1), 75–94. https://doi.org/10.1037/pspi0000110

Durkee, P. K., & Ayers, J. D. (2021). Is facial width-to-height ratio reliably associated with social inferences? *Evolution and Human Behavior*, *42*(6), 583–592. https://doi.org/10.1016/j.evolhumbehav.2021.06.003

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Feinberg, D. R. (2021). Praat Scripts. https://doi.org/10.17605/OSF.IO/HUZ7D

Feinberg, D. R., Jones, B. C., & Armstrong, M. M. (2019). No evidence that men's voice pitch signals formidability. *Trends in Ecology & Evolution*, *34*(3), 190–192. https://doi.org/10.1016/j.tree.2018.12.014

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*(7), 258–267. https://doi.org/10.1016/S1364-6613(00)01494-7

Frederick, D. A., & Peplau, L. A. (2007) The UCLA Body Matrices II: Computer-generated images of men and women varying in body fat and muscularity/breast size to assess body satisfaction and preferences. [Poster presentation]. Annual Meeting of the Society for Personality and Social Psychology, Memphis, TN.

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *Plos One*, *10*(7), e0132726.

Gorris, C., Maccarini, A. R., Vanoni, F., Poggioli, M., Vaschetto, R., Garzaro, M., & Valletti, P. A. (2020). Acoustic analysis of normal voice patterns in Italian adults by using Praat. *Journal of Voice*, *34*(6), 961.e9–961.e18. https://doi.org/10.1016/j.jvoice.2019.04.016

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Han, C., Kandrik, M., Hahn, A. C., Fisher, C. I., Feinberg, D. R., Holzleitner, I. J., & Jones, B. C. (2017). Interrelationships among men's threat potential, facial dominance, and vocal dominance. *Evolutionary Psychology*, *15*(1), 1474704917697332. https://doi.org/10.1177/2F1474704917697332

Hansen, K., Steffens, M. C., Rakić, T., & Wiese, H. (2017). When appearance does not match accent: Neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience*, *12*(3), 507–515. https://doi.org/10.1093/scan/nsw148

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 571–576. https://doi.org/10.1098/rspb.2011.1193

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, *49*(4), 747–754. https://doi.org/10.1016/j.jesp.2013.02.015

Hehman, E., Flake, J. K., & Freeman, J. B. (2015a). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin*, *41*(8), 1123–1134. https://doi.org/10.1177/0146167215591495

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015b). Picking teams: When dominant facial structure is preferred. *Journal of Experimental Social Psychology*, *59*, 51–59. https://doi.org/10.1016/j.jesp.2015.03.007

Hehman, E., Stolier, R. M., Freeman, J. B., Flake, J. K., & Xie, S. Y. (2019). Toward a comprehensive model of face impressions: What we know, what we do not, and paths forward. *Social and Personality Psychology Compass*, *13*(2), e12431. https://doi.org/10.1111/spc3.12431

Jaeger, B., Evans, A. M., & Stel, M., & van Beest (2018). *Who judges a book by its cover?* The prevalence, structure, and correlates of physiognomic beliefs. *PsyArXiv*. https://doi.org/10.31234/osf.io/8dq4x

Jaeger, B., & Jones, A. L. (2020). Which facial features are central in impression formation? *Social Psychological and Personality Science*, *13*(2), 553–561. https://doi.org/10.1177/2F19485506211034979

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Kamachi, M., Hill, H., Lander, K., & Vatikiotis-Bateson, E. (2003). Putting the face to the voice': Matching identity across modality. *Current Biology*, *13*(19), 1709–1714. https://doi.org/10.1016/j.cub.2003.09.005

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698–2704. https://doi.org/10.1098/rspb.2012.0311

Kordsmeyer, T. L., Freund, D., Pita, S. R., Jünger, J., & Penke, L. (2019). Further evidence that facial width-to-height ratio and global facial masculinity are not positively associated with testosterone levels. *Adaptive Human Behavior and Physiology*, *5*(2), 117–130. https://doi.org/10.1007/s40750-018-0105-4

Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science*, *28*(11), 1675–1682. https://doi.org/10.1177/2F0956797617716929

Kramer, R. S. (2017). Sexual dimorphism of facial width-to-height ratio in human skulls and faces: A meta-analytical approach. *Evolution and Human Behavior*, *38*(3), 414–420. https://doi.org/10.1016/j.evolhumbehav.2016.12.002

Krenn, B., & Meier, J. (2018). Does facial width-to-height ratio predict aggressive behavior in association football? *Evolutionary Psychology*, *16*(4), 1–8. https://doi.org/10.1177/2F1474704918818590

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2021). Estimated marginal means, aka least-squares means. [R Package Emmeans Version 1.6. 0]. *Comprehensive R Archive Network*, (CRAN).

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, *34*(4), 273–279. https://doi.org/10.1016/j.evolhumbehav.2013.03.005

Lieberz, K. A., Windmann, S., Geniole, S. N., McCormick, C. M., Mueller-Engelmann, M., Gruener, F., & Steil, R. (2017). The facial width-to-height ratio determines interpersonal distance preferences in the observer. *Aggressive Behavior*, *43*(5), 460–470. https://doi.org/10.1002/ab.21704

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Masi, M., Mattavelli, S., Fasoli, F., & Brambilla, M. (2022). Cross-modal impression updating: Dynamic impression updating from face to voice and the other way around. *British Journal of Social Psychology*, *61*(3), 808–825. https://doi.org/10.1111/bjso.12511

Matsumoto, D., & Hwang, H. C. (2021). Facial width-to-height ratios and deception skill. *Personality and Individual Differences*, *174*, 110683. https://doi.org/10.1016/j.paid.2021.110683

Mavica, L. W., & Barenholtz, E. (2013). Matching voice and face identity from static images. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 307–312. https://doi.org/10.1037/a0030945

McAleer, P., & Belin, P. (2018). The perception of personality traits from voices. In S. Fruholz, & P. Belin (Eds.), *The oxford handbook of voice perception* (pp. 585–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.26

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLoS ONE*, *9*(3), e90779. https://doi.org/10.1371/journal.pone.0090779

McArthur, L. Z., & Baron, R. M. (1983). Toward and ecological theory of social perception. *Psychological Review*, *90*(3), 215–238. https://doi.org/10.1037/0033-295X.90.3.215

Melnikoff, D. E., & Bailey, A. H. (2018). Preferences for moral vs. immoral traits in others are conditional. *Proceedings of the National Academy of Sciences*, *115*(4), E592–E600. https://doi.org/10.1073/pnas.1714945115

Mileva, M., Tompkinson, J., Watt, D., & Burton, A. M. (2018). Audiovisual integration in social evaluation. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 128–138. https://doi.org/10.1037/xhp0000439

Mileva, M., Tompkinson, J., Watt, D., & Burton, A. M. (2020). The role of face and voice cues in predicting the outcome of student representative elections. *Personality and Social Psychology Bulletin*, *46*(4), 617–625. https://doi.org/10.1177/2F0146167219867965

Mileva, V. R., Cowan, M. L., Cobey, K. D., Knowles, K. K., & Little, A. C. (2014). In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences*, *69*, 115–118. https://doi.org/10.1016/j.paid.2014.05.019

O'Connor, J. J., & Barclay, P. (2017). The influence of voice pitch on perceptions of trustworthiness across social contexts. *Evolution and Human Behavior*, *38*(4), 506–512. https://doi.org/10.1016/j.evolhumbehav.2017.03.001

Perugini, M., Gallucci, M., & Costantini, G. (2014). Safeguard power as a protection against imprecise power estimates. *Perspectives on Psychological Science*, *9*(3), 319–332. https://doi.org/10.1177/2F1745691614528519

Pisanski, K., & Bryant, G. A. The evolution of voice perception. In N. S. EidSheim & K. Meizel (Eds.) (2019). The Oxford handbook of voice studies (pp. 269–300). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199982295.013.29

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, *20*(4), 304–318. https://doi.org/10.1016/j.tics.2016.01.002

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, *95*, 89–99. https://doi.org/10.1016/j.anbehav.2014.06.011

Puts, D. A., & Aung, T. (2019). Does men's voice pitch signal formidability? A reply to Feinberg et al. *Trends in Ecology & Evolution*, *34*(3), 189–190. https://doi.org/10.1016/j.tree.2018.12.004

Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of Sex Research*, *49*(2–3), 227–243. https://doi.org/10.1080/00224499.2012.658924

Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., .. & , & Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1829), 20152830. https://doi.org/10.1098/rspb.2015.2830

Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1208–1219. https://doi.org/10.1037/0096-1523.33.5.1208

Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant voices and attractive faces: The contribution of visual and auditory information to integrated person impressions. *Journal of Nonverbal Behavior*, *39*(4), 355–370. https://doi.org/10.1007/s10919-015-0214-8

Schild, C., & Zettler, I. (2021). Linking voice pitch to fighting success in male amateur mixed martial arts athletes and boxers. *Evolutionary Human Sciences*, *3*, e46. https://doi.org/10.1017/ehs.2021.45

Schild, C., Stern, J., & Zettler, I. (2020). Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, *31*(1), 164–175. https://doi.org/10.1093/beheco/arz173

Schroeder, J., & Epley, N. (2016). Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General*, *145*(11), 1427–1437. https://doi.org/10.1037/xge0000214

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–354. https://doi.org/10.1177/0956797610362647

Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2019). Faces tell everything in a just and biologically determined world: Lay theories behind face reading. *Social Psychological and Personality Science*, *10*(1), 62–72. https://doi.org/10.1177/2F1948550617734616

Tatz, J. R., Peynircioğlu, Z. F., & Brent, W. (2020). Face-voice space: Integrating visual and auditory cues in judgments of person distinctiveness. *Attention, Perception, & Psychophysics*, *82*(7), 3710–3727. https://doi.org/10.3758/s13414-020-02084-1

Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, *33*(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*, 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Van Vugt, M. (2009). Sex differences in intergroup competition, aggression, and warfare: The male warrior hypothesis. *Annals of the New York Academy of Sciences*, *1167*(1), 124–134. https://doi.org/10.1111/j.1749-6632.2009.04539.x

Wang, D., Nair, K., Kouchaki, M., Zajac, E. J., & Zhao, X. (2019). A case of evolutionary mismatch? Why facial width-to-height ratio may not predict behavioral tendencies. *Psychological Science*, *30*(7), 1074–1081. https://doi.org/10.1177/2F0956797619849928

Wells, T., Baguley, T., Sergeant, M., & Dunn, A. (2013). Perceptions of human attractiveness comprising face and voice cues. *Archives of Sexual Behavior*, *42*(5), 805–811. https://doi.org/10.1007/s10508-012-0054-0

Weston, E. M., Friday, A. E., & Liò, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *Plos One*, *2*(8), e710. https://doi.org/10.1371/journal.pone.0000710

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of Personality and Social Psychology*, *113*(1), 59–80. https://doi.org/10.1037/pspi0000092

Witkower, Z., & Tracy, J. L. (2019). A facial-action imposter: How head tilt influences perceptions of dominance from a neutral face. *Psychological Science*, *30*(6), 893–906. https://doi.org/10.1177/2F0956797619838762

Young, A. W., Frühholz, S., & Schweinberger, S. R. (2020). Face and voice perception: Understanding commonalities and differences. *Trends in Cognitive Sciences*, *24*(5), 398–410. https://doi.org/10.1016/j.tics.2020.02.001

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/2F0963721416683996

Zhang, J., Hodges-Simeon, C., Gaulin, S. J., & Reid, S. A. (2021). Pitch lowering enhances men's perceived aggressive intent, not fighting ability. *Evolution and Human Behavior*, *42*(1), 51–60. https://doi.org/10.1016/j.evolhumbehav.2020.07.007

Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2015). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior*, *41*(4), 322–330. https://doi.org/10.1002/ab.21544

Zuckerman, M., & Sinicropi, V. (2011). When physical and vocal attractiveness differ: Effects on favorability of interpersonal impressions. *Journal of Nonverbal Behavior*, *35*(2), 75–86. https://doi.org/10.1007/s10919-011-0106-5

How to cite this article: Masi, M., Mattavelli, S., Fasoli, F., & Brambilla, M. (2022). When faces and voices come together: Face width-to-height ratio and voice pitch contribute independently to social perception. *European Journal of Social Psychology*, 1–13. https://doi.org/10.1002/ejsp.2905