

Research



Cite this article: Pinheiro AP, Anikin A, Conde T, Sarzedas J, Chen S, Scott SK, Lima CF. 2021 Emotional authenticity modulates affective and social trait inferences from voices. *Phil. Trans. R. Soc. B* **376**: 20200402. <https://doi.org/10.1098/rstb.2020.0402>

Accepted: 12 May 2021

One contribution of 11 to a theme issue 'Voice modulation: from origin and mechanism to social impact (Part I)'.

Subject Areas:

behaviour, cognition

Keywords:

voice, emotion, authenticity, social traits, acoustics

Authors for correspondence:

Ana P. Pinheiro
e-mail: appinheiro@psicologia.ulisboa.pt
César F. Lima
e-mail: cesar.lima@iscte-iul.pt

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5631307>.

Emotional authenticity modulates affective and social trait inferences from voices

Ana P. Pinheiro¹, Andrey Anikin^{2,3}, Tatiana Conde¹, João Sarzedas¹, Sinead Chen⁴, Sophie K. Scott⁵ and César F. Lima^{5,6}

¹CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal

²Equipe de Neuro-Ethologie Sensorielle (ENES)/Centre de Recherche en Neurosciences de Lyon (CRNL), University of Lyon/Saint-Etienne, CNRS UMR5292, INSERM UMR_S 1028, 42023 Saint-Etienne, France

³Division of Cognitive Science, Lund University, 221 00 Lund, Sweden

⁴National Taiwan University, Taipei City, 10617 Taiwan

⁵Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, UK

⁶Instituto Universitário de Lisboa (ISCTE-IUL), Avenida das Forças Armadas, 1649-026 Lisboa, Portugal

id APP, 0000-0002-7981-3682; AA, 0000-0002-1250-8261; TC, 0000-0002-9223-0924; SKS, 0000-0001-7510-6297; CFL, 0000-0003-3058-7204

The human voice is a primary tool for verbal and nonverbal communication. Studies on laughter emphasize a distinction between spontaneous laughter, which reflects a genuinely felt emotion, and volitional laughter, associated with more intentional communicative acts. Listeners can reliably differentiate the two. It remains unclear, however, if they can detect authenticity in other vocalizations, and whether authenticity determines the affective and social impressions that we form about others. Here, 137 participants listened to laughs and cries that could be spontaneous or volitional and rated them on authenticity, valence, arousal, trustworthiness and dominance. Bayesian mixed models indicated that listeners detect authenticity similarly well in laughter and crying. Speakers were also perceived to be more trustworthy, and in a higher arousal state, when their laughs and cries were spontaneous. Moreover, spontaneous laughs were evaluated as more positive than volitional ones, and we found that the same acoustic features predicted perceived authenticity and trustworthiness in laughter: high pitch, spectral variability and less voicing. For crying, associations between acoustic features and ratings were less reliable. These findings indicate that emotional authenticity shapes affective and social trait inferences from voices, and that the ability to detect authenticity in vocalizations is not limited to laughter.

This article is part of the theme issue 'Voice modulation: from origin and mechanism to social impact (Part I)'.

1. Introduction

The human voice is a rich source of nonverbal information in social interactions. When we listen to someone talking, or laughing with friends, we can rapidly extract cues related to aspects such as the age, sex, identity or emotional state of the speaker. We also form impressions about whether they sound trustworthy or not, or more or less dominant [1]. Most of what we know about voice perception comes from studies using speech stimuli. Examples are the study of emotion perception in speech prosody (e.g. [2–4]) and identity perception in spoken utterances (e.g. [5]). But there is an increasing interest in understanding how we process voices in the absence of concurrent linguistic information, in nonverbal vocalizations such as laughter and crying. Nonverbal vocalizations are distinct from spoken language regarding their underlying articulatory mechanisms [6]. They represent a primitive, universal and efficient form of emotional communication [7–12].

Recent studies highlight that nonverbal vocalizations can vary a lot, consistent with the complexity and variability that characterize vocal signals [13]. They

can vary between speakers, due to differences in the anatomy of the vocal apparatus, for example. They can also vary within the same speaker depending on context. We laugh differently depending on whether we are spontaneously reacting to a funny video, or deliberately trying to show that we agree with our boss in a meeting. A distinction has been made between *spontaneous* and *volitional* vocalizations in several studies on laughter (e.g. [14–17]). Spontaneous laughter is less controlled, reflects a genuinely felt emotion and is typically reactive to outside events. Volitional laughter is part of more flexible and deliberate communicative acts. It is used to convey appreciation, agreement or to deceive others [17]. These two forms of laughter might rely on distinct vocal production mechanisms. Spontaneous laughter has been suggested to be initiated by a complex set of midline structures involved in innate vocalizations (e.g. periaqueductal grey [17,18]). It is characterized by rhythmic respiratory and laryngeal activity, and it typically does not require supralaryngeal articulators [14]. Volitional laughter, on the other hand, might be supported by the same sensorimotor cortical regions that control the production of learned vocalizations such as speech and song (e.g. lateral motor and premotor cortices [17,18]). It involves an increased engagement and fine motor control over breathing and supralaryngeal articulators, similarly to the complex coordination required by speech [14,19].

Spontaneous and volitional laughter also differ acoustically and perceptually. Spontaneous laughter is often higher in pitch, longer in duration, and shows spectral features that differ from volitional laughter; volitional laughter, on the other hand, is more nasal than spontaneous laughter [20]. Perceptually, spontaneous laughter is perceived as more authentic than volitional laughter, showing that listeners can distinguish between the two types of vocalizations (e.g. [20–22]), possibly because authentic laughter works as a highly salient (and honest) signal, effective at automatically capturing attention [23,24]. Sensitivity to laughter authenticity has additionally been shown to be consistent across cultures [25], and to relate to distinct cortical responses. McGettigan *et al.* [26] found that listening to spontaneous laughter elicited increased activation in bilateral superior temporal gyri, whereas listening to volitional laughter elicited increased activation in anterior medial prefrontal and anterior cingulate cortices, suggesting a more active engagement of mentalizing processes when vocalizations are not genuine.

Two important questions remain unanswered. Because most studies probing the authenticity of nonverbal vocalizations are focused on laughter, it remains unclear whether the reported acoustic and perceptual differences extend to other types of vocalizations, such as crying—a biologically salient vocalization that typically signals distress [12,27], is produced from a very early age like laughter [28] and involves complex respiratory, laryngeal and supralaryngeal articulatory activity [29]. Examining authenticity in crying represents a methodological challenge. It is relatively easy to elicit genuine laughter in the laboratory, because laughter is a pervasive emotional expression that can be primed by a diversity of stimuli in social interactions. It is also a behaviourally contagious expression that can be primed solely by another's laughter [17]. The same is not observed for negative vocalizations because they tend to be less contagious and experiencing the associated emotional states is unpleasant, often involving feelings of helplessness and powerlessness [30].

One study that focussed on the perception of a range of positive and negative vocalizations, including crying, found

that listeners categorize vocalizations from YouTube videos of emotional episodes as more authentic than acted vocalizations from published corpora [31]. This is suggestive of a general ability to detect the authenticity of vocalizations, although evidence for vocalizations other than laughter remains scarce and conflicting [32]. Furthermore, using stimuli from online videos is not without problems: the quality of the recordings is often low and not comparable to laboratory recordings; the emotions and their authenticity need to be inferred from contextual cues; it is not possible to have the same speakers across emotion and authenticity conditions; and the fact that speakers are being filmed can affect their expressions. We therefore need to determine whether the ability to distinguish genuine and posed vocalizations extends beyond laughter, using well-controlled stimuli that account for these potential confounds. Being able to detect authenticity is an important social skill in the case of laughter, to avoid deception and to guide decisions to cooperate [14], but it is also an important skill in the case of other vocalizations. For example, volitional crying can be used from early in development (8–12 months) in tactical or manipulative ways to motivate advantageous caregiver attention [28], therefore requiring vigilance on the part of the receiver.

A second underexplored question is whether the authenticity of vocalizations determines how we form affective and social impressions about a speaker. As for affective evaluations, studies on laughter show that spontaneous vocalizations can lead to higher ratings of perceived valence and arousal [20,26]. Speakers are perceived to be in a more positive and aroused state when they laugh spontaneously compared to when they laugh voluntarily. This needs to be replicated and examined for other vocalizations. As for how authenticity modulates social impressions, nothing is known. Social trait evaluations of faces and voices have been proposed to be based on two core dimensions: trustworthiness and dominance [1,33]. We routinely evaluate whether someone looks or sounds trustworthy or dominant rapidly, often within milliseconds. These judgements are hardly grounded in truth (their accuracy is low), and they are thought to reflect an overgeneralization effect—we generalize to infer that someone has a stable trait (e.g. trustworthiness), merely because their momentary facial or vocal cues (e.g. a smile) resemble expressions that we associate with that trait [34–36]. Nevertheless, such judgements have been shown to be relatively consistent across raters, and to affect our decisions, attitudes and behaviours [37,38]. Studies on facial expressions have shown that people producing Duchenne smiles, which include activation of the muscle that causes wrinkles around the eyes and are associated with genuine happiness, are evaluated as more trustworthy than those producing non-Duchenne smiles [39]. Effects of the Duchenne marker on dominance are less clear [40]. It is plausible that a speaker might be perceived as more trustworthy when their vocalizations are spontaneous compared to when they are volitional.

In the current study, we examined whether listeners detect the authenticity of laughter and crying vocalizations, using well-controlled crying stimuli generated via emotion induction in a laboratory setting. Based on previous findings [20–22,31], we predicted that spontaneous vocalizations would be associated with higher authenticity ratings. We also asked whether both *objective* (stimulus-based) and *perceived* emotional authenticity shape how listeners evaluate the affective state of the speaker, namely valence and arousal, as well as

their trustworthiness and dominance, as predicted by the over-generalization hypothesis [34,35]. Spontaneous vocalizations were expected to be rated higher in arousal and more extreme in valence: more positive in the case of spontaneous laughter, and more negative in the case of spontaneous crying. Based on findings from smile authenticity [39], spontaneous vocalizations were also expected to produce higher perceptions of trustworthiness. As for potential effects on impressions of dominance, our approach was exploratory. Finally, as additional exploratory questions, we examined acoustic differences between spontaneous and volitional vocalizations, and how acoustic features predicted subjective ratings. We wanted to contribute to the still scarce literature on the acoustic correlates of emotional authenticity in the voice [20,26] and were interested in exploring the extent to which the acoustic features that signal authenticity match those that also signal other affective and social inferences.

2. Methods

(a) Participants

One hundred and thirty-seven volunteers participated in the study ($M_{\text{age}} = 21.64$, $s.d. = 6.13$, range = 19–57 years; 115 female). All were European Portuguese native speakers and had normal hearing and normal or corrected-to-normal visual acuity.

The study was approved by the ethics committee of the Faculty of Psychology—University of Lisbon. Before taking part, all participants were informed about the procedures and provided written informed consent. They received course credit for their participation.

We used Bayesian inference in our analyses, which relies on estimates of uncertainty and not on p -values. Nevertheless, our sample size can also be considered appropriate according to the standard null hypothesis significance testing approach. An *a priori* power analysis with G*Power 3.1 [41] indicated that a sample size of at least 84 would be required to detect significant correlations of $r = 0.30$ or larger between variables, considering an alpha level of 0.05 and a power of 0.80.

(b) Stimuli

The experimental stimuli consisted of 75 vocalizations, divided into four conditions: 19 spontaneous laughs, 19 volitional laughs, 19 spontaneous cries and 18 volitional cries. They were selected from a larger set of stimuli recorded by six speakers (three women) within a sound-proof anechoic chamber at University College London. The speakers were young and middle-aged adults, with ages between 24 and 48 years. They were not actors, but all had some experience of recording vocal materials (e.g. because they had taken part in similar recording sessions previously). They also indicated beforehand that they felt they would be able to produce both volitional and spontaneous laughter and crying.

This set of laughter vocalizations has been used in previous behavioural and neuroimaging experiments focused on authenticity detection [20–22,42]. Crying vocalizations have also been used in prior studies [22,43], but this is the first one to address how they are perceived regarding their authenticity.

To record volitional laughter and crying, the six speakers were asked to intentionally produce these vocal expressions in the absence of a corresponding emotional eliciting event, and to make them sound as natural and credible as possible. This is in line with the procedure typically used for the recording of acted stimuli [44–47].

As for genuine vocalizations, spontaneous laughter was elicited using an amusement induction procedure in a social interactive setting: speakers watched video clips that they had

previously identified as amusing and that would easily make them laugh. The experimenters knew the speakers well and interacted with them during the recording session to promote the naturalness and the social nature of the laughs. Spontaneous crying was also obtained via an emotion induction procedure. Speakers were asked to recall difficult (upsetting) past episodes and/or to initially produce volitional crying to promote a transition into spontaneous crying reflecting a genuine experience of sadness. All speakers confirmed that they were able to cry spontaneously and were asked to indicate the point in the recording that marked the onset/transition to spontaneous crying, even though that was perceptually clear to the experimenters in most instances. During the debriefing, all speakers reported having experienced feelings of amusement and sadness throughout and after recording the corresponding spontaneous expressions. They also reported that they felt much less control over their vocalizations when they were spontaneous, both in the case of laughter and crying, compared to when they were volitional.

(c) Procedure

The experiment was conducted in a quiet room in a laboratory setting at the Faculty of Psychology, University of Lisbon. Participants were tested in small group sessions with up to eight participants per session.

Vocal stimuli were presented via headphones, and stimulus presentation and response recording were controlled using *Qualtrics* software (see <https://www.qualtrics.com>). Participants were instructed to rate the sounds as quickly as possible, following their first impressions. After the presentation of each vocalization, participants were first asked to indicate the emotion that best characterized the sound in a three-alternative forced-choice categorization: ‘sadness’, ‘neutral’ or ‘happiness’. They then rated the vocalization regarding the dimensions of emotional authenticity, valence and arousal, as well as the social traits of trustworthiness and dominance of the speaker. Nine-point scales were used, from 1 (minimum) to 9 (maximum). The emotion categorization task was included as a manipulation check, i.e. to confirm that the stimuli conveyed the emotions they were expected to, and that the main findings (focused on the affective and social ratings) could not be explained by difficulties with perceiving those emotions.

Vocalizations were presented once, in a pseudo-random order to avoid the presentation of more than two consecutive vocalizations from the same category. Before the experiment, two practice examples were provided: a crying and a laughter exemplar from the Montreal Affective Voices [44]. The session lasted around 45 min.

(d) Data analysis

(i) Behavioural data

Statistical analyses were performed on unaggregated responses from individual trials using Bayesian mixed models and the *brms* R package [48]. All results were summarized as the medians of posterior distributions and 95% credible intervals (CI). When contrasting two conditions (e.g. spontaneous versus volitional vocalizations), the CI includes the most credible values for the difference given the data and the model and, if that does not include 0, we can infer that there is evidence in favour of an actual difference between conditions (see electronic supplementary material). The code used for data analysis and the full dataset can be found here: https://osf.io/57syv/?view_only=c98e91f70a2d49e8ad902bbde2a4482c.

(ii) Acoustic features

The audio files were downsampled to 22050 Hz, high-passed over 90 Hz to remove low-frequency noise and analysed acoustically with the *soundgen* R package [49]. Intonation contours were

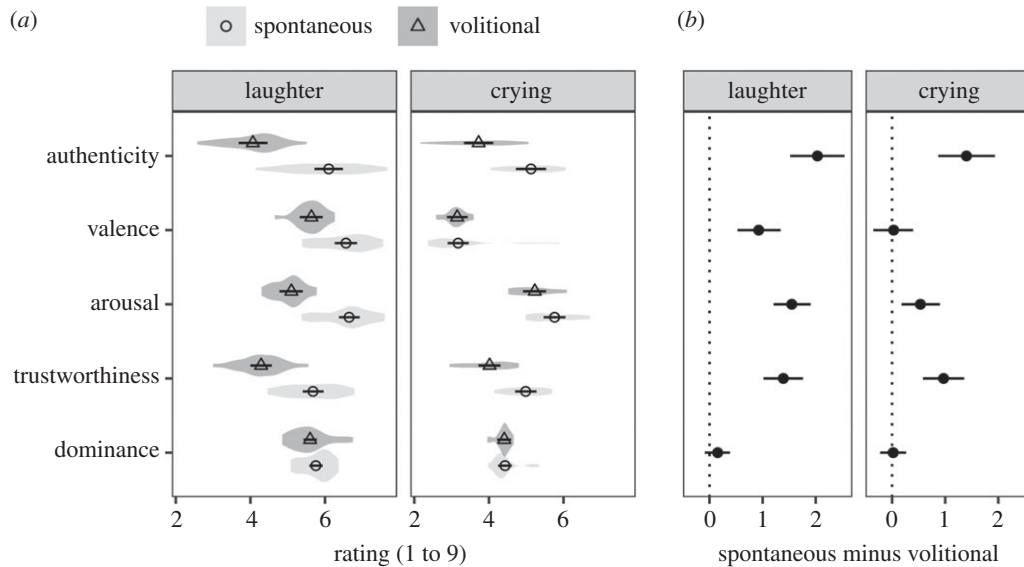


Figure 1. Ratings on each of five scales as a function of the spontaneous or volitional nature of the rated vocalizations (a) and the difference in ratings between spontaneous and volitional vocalizations (b). The points and error bars show fitted values (medians of posterior distribution with 95% CI), and violin plots show the distribution of mean observed values per stimulus.

manually verified and, if necessary, corrected using the *pitch_app()* interactive environment. When appropriate, acoustic descriptives were summarized as the median value and standard deviation (SD) across the entire sound duration. Among the potentially large number of quantifiable acoustic characteristics, we focused on nine key variables, chosen *a priori* because they are theoretically meaningful, have often been reported in earlier studies and can be measured reliably (e.g. [31]):

- Duration (in seconds [s]): the duration of a stimulus without counting silent frames at the beginning and end;
- Harmonics-to-noise ratio (HNR), median and SD (in decibels [dB]): a measure of pitch quality or tonality calculated only for voiced frames;
- Novelty: a measure of spectral variability derived from the self-similarity matrix (SSM) of a vocalization by sliding a 200 ms Gaussian checkerboard matrix along the SSM's diagonal;
- Pitch, median and SD (in hertz [Hz]): manually verified fundamental frequency or perceived tone height;
- Spectral centroid, median and SD (Hz): the first spectral moment or centre of gravity of the spectrum of voiced frames, which perceptually corresponds to timbral brightness;
- Voiced (in percentage [%]): the proportion of voiced frames.

The variables measured in Hz were transformed to the more perceptually relevant logarithmic scale, following which all variables were scaled to have a mean of 0 and SD of 1. We used median rather than mean values because medians are more robust to outliers, such as frames with incorrectly measured pitch or external noise.

To test the effect of acoustic predictors on the ratings, we again used multivariate ordinal regression and predicted the ratings on all five scales as a function of the nine measured acoustic features (see electronic supplementary material).

3. Results

(a) Affective and social ratings of spontaneous and volitional vocalizations

As a measure of inter-rater agreement, we aggregated the ratings of each sound on each of the five scales and calculated

the mean Pearson's correlation between the responses of each participant and these aggregated ratings. Within each vocalization type (laughter and crying), correlations ranged from 0.44 to 0.62 for four scales: authenticity, valence, arousal and trustworthiness. They were considerably lower for the dominance scale (0.37 for laughter and 0.20 for crying). Likewise, the intraclass correlation coefficient, estimated using a two-way random model and absolute agreement, revealed lower reliability for the dominance scale (less than 0.1) compared to the other four scales (0.1–0.3). As dominance ratings were less consistent, the results for this scale should be treated with caution.

The accuracy of emotion recognition in the forced-choice classification task was above 95% (see electronic supplementary material), confirming that participants accurately recognized the conveyed emotions as expected. Figure 1 shows how spontaneous and volitional laughter and crying were rated on all five scales. First, we tested the hypothesis that listeners can detect the authenticity of laughter and crying vocalizations. In line with our prediction, spontaneous vocalizations were rated as 1.72 points more authentic than volitional ones (95% CI [1.35, 2.09]). The difference was 2.03 points for laughter (95% CI [1.52, 2.54]) and 1.40 points for crying (95% CI [0.87, 1.94]). The difference between laughter and crying in terms of this authenticity contrast was not statistically robust (0.63 points higher for laughter, 95% CI [−0.13, 1.36]). Laughter was overall judged to be slightly more authentic than crying (0.66 points, 95% CI [0.29, 1.03]).

If we consider that a perceived authenticity rating is 'correct' when it is above the midpoint of the 1–9 scale (greater than 5), and below in the case of volitional vocalizations, the overall accuracy of recognizing authenticity was 60.6% (95% CI [55.6, 65.7]). This varied across vocalization types: 65.8% for spontaneous laughter (95% CI [55.2, 75.6]), 62.3% for volitional laughter (95% CI [51.3, 71.8]), 41.9% for spontaneous crying (95% CI [31.6, 52.9]) and 72.6% for volitional crying (95% CI [62.7, 81.4]). Despite the slight bias to treat laughter as more authentic than crying, there were no statistically robust differences in accuracy of authenticity detection when comparing volitional with spontaneous stimuli (13.4% higher accuracy

for volitional, 95% CI [−2.8, 29.6]), or laughs with cries (6.7% higher accuracy for laughs, 95% CI [−9.5, 22.8]).

Then, we tested the hypothesis that objective stimulus authenticity modulates affective and social evaluations of listeners. Regarding valence ratings, laughter was generally rated as more positive (6.10, 95% CI [5.88, 6.32]) than crying (3.16, 95% CI [2.96, 3.37]), as expected (figure 1). Objective authenticity also played a role in the case of laughter. Spontaneous laughs were rated as 0.93 points more positive than volitional laughs (95% CI [0.53, 1.34]). No such effect was found for crying, for which the difference was only 0.03 points (95% CI [−0.36, 0.4]). In general, spontaneous vocalizations were also rated as 1.04 points (95% CI [0.8, 1.29]) more arousing compared to volitional vocalizations, by 1.55 points for laughs (95% CI [1.21, 1.91]) and 0.53 for cries (95% CI [0.18, 0.9]), specifically. Averaging across spontaneous and volitional stimuli, laughter was rated as 0.38 points (95% CI [0.13, 0.62]) higher on arousal compared to crying.

As for trait inferences, spontaneous expressions were perceived as more trustworthy. The difference in the ratings was 1.39 points for laughter (95% CI [1.02, 1.76]) and 0.97 for crying (95% CI [0.58, 1.36]; see figure 1*b*). The effect of authenticity on trustworthiness ratings was similar for laughter and crying (0.42 higher for laughter, 95% CI [−0.13, 0.95]). Averaging across spontaneous and volitional vocalizations, trustworthiness ratings were 0.48 points higher for laughter compared to crying (95% CI [0.22, 0.75]). No differences were found for dominance ratings: inferences were similar for spontaneous and volitional vocalizations (a difference of 0.15 points for laughter, 95% CI [−0.09, 0.39] and of 0.02 points for crying, 95% CI [−0.23, 0.27]).

In sum, in line with our predictions, spontaneous vocalizations were rated as more authentic than volitional ones. Spontaneous (versus volitional) vocalizations were also perceived as more arousing and trustworthy. Objective authenticity also affected the perceived valence of the voice: spontaneous laughs were rated as more positive than their volitional counterparts. However, objective authenticity did not affect valence perception for crying, or inferences of dominance for both laughter and crying.

(b) Predicting social inferences from affective ratings

We also tested the hypothesis that perceived authenticity (in addition to objective authenticity) predicted social trait inferences. First, we examined the correlations among the five rating scales and found that they were small-to-moderate (all r s < 0.3 for crying and < 0.65 for laughter; see electronic supplementary material for a full correlation table). However, there was a strong linear relationship between perceived authenticity and trustworthiness ratings, $r = 0.83$. Thus, both objective and perceived authenticity were associated with inferences of higher trustworthiness. Notably, correlations between perceived authenticity and arousal, a variable that has been highlighted as a potential marker of authenticity [16,20], were much lower both for laughter ($r = 0.42$) and for crying ($r = 0.27$).

To model the relationship between the five scales in more detail, we predicted social ratings (trustworthiness and dominance) from authenticity, valence and arousal ratings (we were primarily interested in authenticity; valence and arousal ratings were included for completeness). The

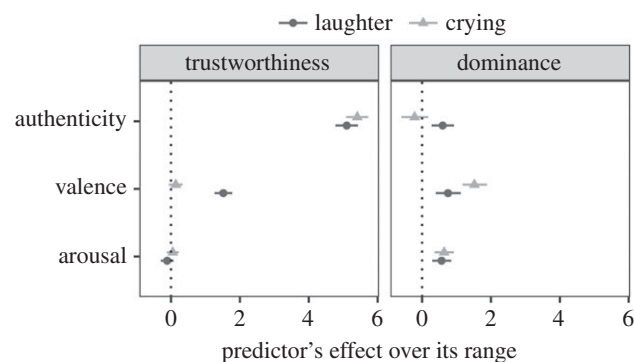


Figure 2. Effects of authenticity, valence and arousal ratings on the perceived trustworthiness and dominance of the speaker, on a 1–9 scale in multiple regression. Medians of posterior distributions and 95% CIs.

analysis revealed that authenticity ratings strongly predicted trustworthiness ratings for laughter (5.10 points higher trustworthiness for an increase in authenticity ratings from 1 to 9, 95% CI [4.78, 5.42]) and crying (5.41, 95% CI [5.09, 5.71]; figure 2). Valence ratings also predicted trustworthiness, but the effect was much smaller, and limited to laughter (1.51 points, 95% CI [1.27, 1.77]); for crying, 0.14 points, 95% CI [−0.06, 0.34]). Arousal ratings were not credibly related to trustworthiness in either laughter (−0.11 points, 95% CI [−0.3, 0.08]) or crying (0.05 points, 95% CI [−0.12, 0.22]).

Perceived authenticity in laughter predicted evaluations of dominance, albeit weakly (0.6 points, 95% CI [0.26, 0.93]), but not in crying (−0.21 points, 95% CI [−0.62, 0.17]). Perceived arousal and valence in both laughter and crying also predicted perceived dominance (effect sizes approx. 1 point on a 1–9 scale, figure 2).

In sum, we confirmed that perceived authenticity predicted social trait inferences, particularly in the case of trustworthiness—laughs and cries perceived as more spontaneous were also perceived as more trustworthy. By contrast, perceived valence and arousal had no or little effects on trustworthiness. The effects of authenticity and affective ratings on inferences of dominance were generally small.

(c) Predicting affective and social ratings from acoustic features of vocalizations

Finally, we examined acoustic differences between spontaneous and volitional vocalizations, and whether acoustic features predict listeners' subjective ratings. As described in §2, we focused on nine theoretically meaningful acoustic characteristics of nonverbal vocalizations. Spontaneous laughs differed from volitional ones in six of these measures (see electronic supplementary material, figure S1). Spontaneous laughs had higher (1.66 SD, 95% CI [1.13, 2.21]) and more variable (1.06, 95% CI [0.47, 1.67]) fundamental frequency or pitch; brighter timbre (0.93, 95% CI [0.41, 1.43] higher spectral centroid); higher (0.68, 95% CI [0.15, 1.21]) and more variable (0.92, 95% CI [0.35, 1.51]) HNR in voiced frames; and greater general variability (0.51, 95% CI [0.01, 1.01] higher novelty). Spontaneous and volitional cries were slightly less distinct acoustically. Spontaneous cries had slightly higher pitch (0.41 SD, 95% CI [0.02, 0.81]); more variable timbral brightness (0.68, 95% CI [0.12, 1.23] higher SD of spectral centroid); and less voicing (−0.79, 95% CI [−1.48, −0.11]). Because the number of spontaneous and volitional

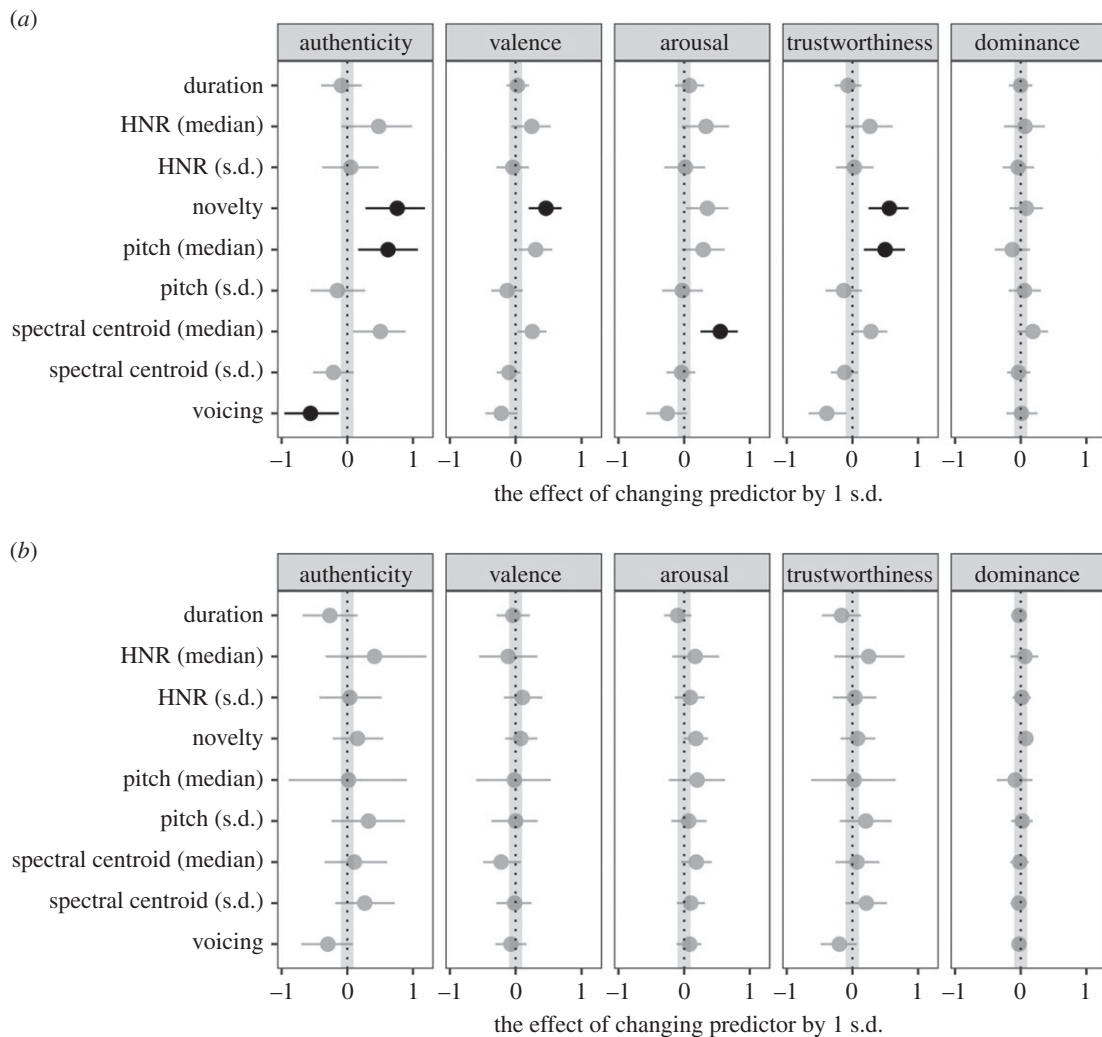


Figure 3. Predicting ratings on the five scales of laughter (a) and crying (b) from acoustic characteristics: medians of posterior distribution and 95% CIs from multivariate ordinal regression. The highlighted effects (in black) correspond to the ones that clear the region of practical equivalence (ROPE) of $(-0.1, 0.1)$, indicating an effect of at least 0.1 on the rating scale (1 to 9) when changing the predictor by 1 SD. HNR = harmonics-to-noise ratio.

vocalizations of each type is relatively small ($n = 18\text{--}19$ per category), this analysis should be seen as descriptive.

We then tested how acoustic features affected the ratings on each of the five scales, separately for laughter and crying (figure 3). Because many of the acoustic predictors are correlated, we estimated their partial effects in multiple regressions. High-pitched and generally variable laughs were judged to be more authentic and trustworthy. This was indicated by the positive effects of median pitch and novelty on authenticity and trustworthiness ratings: 0.62 (95% CI [0.16, 1.08]) and 0.5 (95% CI [0.17, 0.8]) increase in authenticity and trustworthiness ratings, respectively, for a 1 SD increase in pitch; and 0.76 (95% CI [0.28, 1.18]) and 0.56 (95% CI [0.24, 0.85]) increase in perceived authenticity and trustworthiness, respectively, for a 1 SD increase in novelty. Laughs with a smaller proportion of voiced frames were also judged to be more authentic (0.56, 95% CI [0.13, 0.96]).

A shift of spectral energy towards higher harmonics, indicative of increased vocal effort and a bright voice (spectral centroid), predicted higher perceived arousal (0.55, 95% CI [0.25, 0.82]). It also predicted higher authenticity, although the effect did not clear the region of practical equivalence (ROPE; 0.50, 95% CI [0.09, 0.89]). More positive valence ratings of laughs were primarily predicted by greater novelty (0.46, 95% CI [0.2, 0.7]), although there were also statistically

uncertain effects of pitch (0.31, 95% CI [0.05, 0.56]) and spectral centroid (0.25, 95% CI [0.03, 0.47]).

In sum, laughs were perceived to be more authentic and trustworthy when they were high-pitched and very variable. Increased vocal effort signalled high arousal and to some extent perceived authenticity as well. We did not observe any statistically robust acoustic predictors of dominance ratings in laughs, or of any ratings in cries.

4. Discussion

Our findings confirm that the acoustic and perceptual differences between spontaneous and volitional vocalizations extend to vocalizations of negative valence. We observed that the generation of laughs and cries in spontaneous and volitional contexts was effective in leading to changes in perceived authenticity. That is, listeners could reliably judge spontaneous laughs and cries as being more authentic than their volitional counterparts. Moreover, the magnitude of the authenticity distinction in the ratings was similar across the two types of vocalizations. This capacity to detect the authenticity of spontaneous and volitional vocal expressions might confer advantages in social interactions, for example to avoid deception [14,40], and is consistent with the notion

that authentic vocalizations might function as a highly salient signal that automatically captures attention [23,24].

The authenticity of vocalizations shaped the perception of the affective qualities of vocalizations, in good agreement with previous studies on laughter (e.g. [14,20,26]). Specifically, spontaneous laughs (but not cries) were perceived as more pleasant than their volitional counterparts. Both spontaneous laughs and cries were perceived as more arousing (see also [20]). We also showed that objective (stimulus-based) and perceived authenticity affects how listeners form social impressions about a speaker. Social trait inference has been found to interact with emotional perception [1,50–52]. For example, we often rely on transient signals (e.g. the emotional quality of the voice) to make inferences of more stable characteristics of a speaker, such as whether they are trustworthy or friendly (overgeneralization hypothesis [34,35]). We extend these findings by showing that authenticity strongly predicts how trustworthy a voice is perceived to be, irrespective of vocalization type. Laughs and cries produced spontaneously, and perceived as more authentic, were also evaluated as more trustworthy.

Of note, the weight of affective cues in social trait evaluation differed for trustworthiness and dominance. Trustworthiness was strongly predicted by perceived authenticity, less by valence (and only in laughter), and *not* by arousal. These findings document for the first time that authenticity might be signalling something unique that arousal alone does not. Authenticity often correlates with arousal [16,20], as increased arousal has been linked to the presence of ‘hard-to-fake’ properties of voices, but our findings suggest they reflect dissociable dimensions of vocalizations. Supporting this notion, studies using facial expressions indicate that spontaneous and acted smiles are accurately discriminated even when the stimuli are matched for perceived arousal [53]. The links between pleasantness (valence) and trustworthiness of vocalizations confirm the alignment of the two variables in the two-dimensional social voice space [1,54], as they are both related to approachability in social interactions. By contrast, prediction of social inference from affective ratings was generally less robust in the case of dominance, which could be related to the lower agreement among participants when judging how dominant a speaker sounded. Alternatively, these findings may highlight the primary role of trustworthiness (relative to dominance) in social evaluations [55,56].

However, the effects of objective authenticity on affective and social ratings differed for laughs and cries. Spontaneous laughs were perceived as more positive than volitional laughs, but spontaneous cries were not perceived as more negative than their volitional counterparts. Additionally, spontaneous cries were associated with the lowest authenticity recognition accuracy in our derived measure of accuracy (possibly as a result of the lower authenticity ratings provided for crying compared to laughter in general). Furthermore, differences in the perceived arousal of spontaneous versus volitional vocalizations were smaller for cries than for laughs. These findings may reflect a relative difficulty in producing spontaneous crying in an experimental context, which might have made the authenticity of cries relatively less salient and recognizable. Alternatively, it should also be noted that both spontaneous and volitional laughs are highly prevalent expressions in daily social interactions [17,56], while crying is expressed much less often, particularly by adults (compared to children [57–59]). It is therefore plausible that, based on

differences in exposure (and use), listeners find it relatively easier to evaluate authenticity in laughter compared to crying.

Our study also identified which acoustic characteristics predict affective and social judgements of speakers. Authenticity in laughs was predicted by more high-frequency energy (i.e. higher pitch and spectral slope) and variability (novelty). There was considerable overlap in the acoustic features predicting authenticity and trustworthiness ratings in laughs, consistent with the strong correlation between the two rating scales. This suggests that participants might be partly using the same acoustic ‘code’ to make inferences about these two aspects of laughter. The link between authenticity and trustworthiness might also be related to the social function of laughter, namely in establishing and maintaining social bonds [17,60,61], which may occur via increased trust in others. In previous studies, higher pitch has been shown to strongly affect social inference and, specifically, to be associated with higher trustworthiness (e.g. [62]), namely mitigating the aversiveness of spoken words with antisocial content (e.g. ‘cheater’, ‘corrupt’ [63]). Higher pitch and more variable acoustic parameters have also been previously associated with trustworthiness/valence [1,62] and with increased trusting behaviours toward the speaker [64].

Affective and social ratings of cries could not be related to the acoustic features tested here. This suggests that the acoustic hallmarks of authenticity partly differ for positive (laughs) and negative (cries) nonverbal vocalizations. Since our analysis only focused on nine acoustic features, one possibility is that these specific cues play a more subtle role when predicting affective and social evaluations of cries. This could also explain the smaller difference in perceived authenticity between spontaneous and volitional crying compared to laughter. Additionally, as noted before, the particularly challenging task of producing spontaneous crying in an experimental setting could have accounted for the smaller acoustic differences between spontaneous and volitional cries. Studies testing other acoustic parameters are therefore warranted.

Limitations of the current study include the relatively small number of stimuli per condition and also the fact that vocalizations were pre-selected from a larger set, i.e. we only included part of the recorded stimuli. The current findings should be replicated in future studies that include a larger number of stimuli and other vocal emotions (e.g. negative vocalizations such as anger). Another aspect is that we only obtained information about the speaker’s affective states in an informal way, at the debriefing stage—in future attempts to record spontaneous vocalizations, it will be important to address this issue in a more systematic and quantitative way.

5. Conclusion

The present study provides evidence that listeners can reliably infer the emotional authenticity of laughter and crying sounds. It also indicates that emotional authenticity shapes how listeners evaluate the affective state of a speaker, in terms of valence and arousal, as well as how they make social trait inferences, namely regarding trustworthiness. We provide the first demonstration that spontaneous vocal expressions are perceived to be more trustworthy. Moreover, we show that spontaneous vocal expressions differ from volitional ones in several acoustic features, and that the constellation of acoustic differences is partly unique for laughter and crying. It was difficult to predict

the ratings of cries from acoustic features, but for laughter, we could see that the acoustic predictors of perceived authenticity and trustworthiness are similar. High-pitched and acoustically variable laughs were considered to be both more genuine and more trustworthy.

Our findings have implications for theories of social perception (e.g. [65]). They indicate that authenticity should be considered when accounting for how listeners form social impressions from voices. They raise the interesting possibility that genuine vocal expressions may lead to more trusting, cooperative and prosocial behaviour in social interactions, a hypothesis that needs to be addressed in future studies.

Ethics. The study was approved by the ethics committee of Faculty of Psychology University of Lisbon. Before taking part, all participants were informed about the procedures and provided written informed consent.

References

- McAlear P, Todorov A, Belin P. 2014 How do you say 'hello'? Personality impressions from brief novel voices. *PLoS ONE* **9**, e90779. (doi:10.1371/journal.pone.0090779)
- Schirmer A, Kotz SA. 2006 Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* **10**, 24–30. (doi:10.1016/j.tics.2005.11.009)
- Grandjean D. 2021 Brain networks of emotional prosody processing. *Emot. Rev.* **13**, 34–43. (doi:10.1177/1754073919898522)
- Pinheiro AP, Del Re E, Mezin J, Nestor PG, Rauber A, McCarley RW, Gonçalves OF, Niznikiewicz MA. 2013 Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychol. Med.* **43**, 603–618. (doi:10.1017/S003329171200133X)
- Lavan N, Burston LFK, Garrido L. 2019 How many voices did you hear? Natural variability disrupts identity perception from unfamiliar voices. *Br. J. Psychol.* **110**, 576–593. (doi:10.1111/bjop.12348)
- Scott SK, Sauter D, McGettigan C. 2010 Brain mechanisms for processing perceived emotional vocalizations in humans. In *Handbook of behavioral neuroscience* (ed. SM Brudzynski), pp. 187–197. London, UK: Elsevier. (doi:10.1016/B978-0-12-374593-4.00019-X)
- Castiájo P, Pinheiro AP. 2019 Decoding emotions from nonverbal vocalizations: how much voice signal is enough? *Motiv. Emot.* **43**, 803–813. (doi:10.1007/s11031-019-09783-9)
- Lima CF, Anikin A, Monteiro AC, Scott SK, Castro SL. 2018 Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion* **19**, 219–233. (doi:10.1037/emo0000429)
- Pinheiro AP, Lima D, Albuquerque PB, Anikin A, Lima CF. 2019 Spatial location and emotion modulate voice perception. *Cogn. Emot.* **33**, 1577–1586. (doi:10.1080/02699931.2019.1586647)
- Sauter DA, Eisner F, Calder AJ, Scott SK. 2010 Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* **63**, 2251–2272. (doi:10.1080/17470211003721642)
- Sauter DA, Eisner F, Ekman P, Scott SK. 2010 Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl Acad. Sci. USA* **107**, 2408–2412. (doi:10.1073/pnas.0908239106)
- Sauter DA, Crasborn O, Engels T, Kamiloglu RG, Sun R, Eisner F, Haun DBM. 2020 Human emotional vocalizations can develop in the absence of auditory learning. *Emotion* **20**, 1435–1445. (doi:10.1037/emo0000654)
- Lavan N, Burton AM, Scott SK, McGettigan C. 2019 Flexible voices: identity perception from variable vocal signals. *Psychon. Bull. Rev.* **26**, 90–102. (doi:10.3758/s13423-018-1497-7)
- Bryant GA, Aktipis CA. 2014 The animal nature of spontaneous human laughter. *Evol. Hum. Behav.* **35**, 327–335. (doi:10.1016/j.evolhumbehav.2014.03.003)
- Gervais M, Wilson DS. 2005 The evolution and functions of laughter and humor: a synthetic approach. *Q. Rev. Biol.* **80**, 395–430. (doi:10.1086/498281)
- McKeown G, Sneddon I, Curran W. 2015 Gender differences in the perceptions of genuine and simulated laughter and amused facial expressions. *Emot. Rev.* **7**, 30–38. (doi:10.1177/1754073914544475)
- Scott SK, Lavan N, Chen S, McGettigan C. 2014 The social life of laughter. *Trends Cogn. Sci.* **18**, 618–620. (doi:10.1016/j.tics.2014.09.002)
- Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016 Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* **20**, 304–318. (doi:10.1016/j.tics.2016.01.002)
- Bryant GA. 2020 The evolution of human vocal emotion. *Emot. Rev.* **13**, 25–33. (doi:10.1177/1754073920930791)
- Lavan N, Scott SK, McGettigan C. 2016 Laugh like you mean it: authenticity modulates acoustic, physiological and perceptual properties of laughter. *J. Nonverbal Behav.* **40**, 133–149. (doi:10.1007/s10919-015-0222-8)
- Neves L, Cordeiro C, Scott SK, Castro SL, Lima CF. 2018 High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. *Q. J. Exp. Psychol.* **71**, 2355–2363. (doi:10.1177/1747021817741800)
- O'Nions E, Lima CF, Scott SK, Roberts R, McCrory EJ, Viding E. 2017 Reduced laughter contagion in boys at risk for psychopathy. *Curr. Biol.* **27**, 3049–3055. (doi:10.1016/j.cub.2017.08.062)
- Ohman A, Flykt A, Esteves F, Öhman A, Flykt A, Esteves F. 2001 Emotion drives attention: detecting the snake in the grass. *J. Exp. Psychol. Gen.* **130**, 466–478. (doi:10.1037/0096-3445.130.3.466)
- Pinheiro AP, Barros C, Dias M, Kotz SA. 2017 Laughter catches attention! *Biol. Psychol.* **130**, 11–21. (doi:10.1016/j.biopsycho.2017.09.012)
- Bryant GA *et al.* 2018 The perception of spontaneous and volitional laughter across 21 societies. *Psychol. Sci.* **29**, 1515–1525. (doi:10.1177/0956797618778235)
- McGettigan C, Walsh E, Jessop R, Agnew ZK, Sauter DA, Warren JE, Scott SK. 2015 Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb. Cortex* **25**, 246–257. (doi:10.1093/cercor/bht227)
- Blasi A *et al.* 2011 Early specialization for voice and emotion processing in the infant brain. *Curr. Biol.* **21**, 1220–1224. (doi:10.1016/j.cub.2011.06.009)
- Nakayama H. 2010 Development of infant crying behavior: a longitudinal case study. *Infant Behav. Dev.* **33**, 463–471. (doi:10.1016/j.infbeh.2010.05.002)
- Bylsma LM, Gračanin A, Vingerhoets AJJM. 2019 The neurobiology of human crying. *Clin. Auton. Res.* **29**, 63–73. (doi:10.1007/s10286-018-0526-y)

30. Vingerhoets AJJM, Cornelius RR, Van Heck GL, Becht MC. 2000 Adult crying: a model and review of the literature. *Rev. Gen. Psychol.* **4**, 354–377. (doi:10.1037/1089-2680.4.4.354)
31. Anikin A, Lima CF. 2018 Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Q. J. Exp. Psychol.* **71**, 622–641. (doi:10.1080/17470218.2016.1270976)
32. Atias D, Aviezer H. 2020 Real-life and posed vocalizations to lottery wins differ fundamentally in their perceived valence. *Emotion* (doi:10.1037/emo0000931)
33. Oosterhof NN, Todorov A. 2008 The functional basis of face evaluation. *Proc. Natl Acad. Sci. USA* **105**, 11 087–11 092. (doi:10.1073/pnas.0805664105)
34. McArthur LZ, Baron RM. 1983 Toward an ecological theory of social perception. *Psychol. Rev.* **90**, 215–238. (doi:10.1037/0033-295x.90.3.215)
35. Zebrowitz LA, Collins MA. 1997 Accurate social perception at zero acquaintance: the affordances of a Gibsonian approach. *Personal. Soc. Psychol. Rev.* **1**, 204–223. (doi:10.1207/s15327957pspr0103_2)
36. Zebrowitz LA, Montepare JM. 2008 Social psychological face perception: why appearance matters. *Soc. Personal. Psychol. Compass* **2**, 1497–1517. (doi:10.1111/j.1751-9004.2008.00109.x)
37. Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P. 2015 Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu. Rev. Psychol.* **66**, 519–545. (doi:10.1146/annurev-psych-113011-143831)
38. Todorov A, Said CP, Engell AD, Oosterhof NN. 2008 Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* **12**, 455–460. (doi:10.1016/j.tics.2008.10.001)
39. Gunnery SD, Ruben MA. 2016 Perceptions of Duchenne and non-Duchenne smiles: a meta-analysis. *Cogn. Emot.* **30**, 501–515. (doi:10.1080/02699931.2015.1018817)
40. Quadflieg S, Vermeulen N, Rossion B. 2013 Differential reliance on the Duchenne marker during smile evaluations and person judgments. *J. Nonverbal Behav.* **37**, 69–77. (doi:10.1007/s10919-013-0147-z)
41. Faul F, Erdfelder E, Lang AG, Buchner A. 2007 G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191. (doi:10.3758/BF03193146)
42. Lima CF, Brancatisano O, Fancourt A, Müllensiefen D, Scott SK, Warren JD, Stewart L. 2016 Impaired socio-emotional processing in a developmental music disorder. *Sci. Rep.* **6**, 34911. (doi:10.1038/srep34911)
43. Lavan N, Lima CF, Harvey H, Scott SK, McGettigan C. 2015 I thought that I heard you laughing: contextual facial expressions modulate the perception of authentic laughter and crying. *Cogn. Emot.* **29**, 935–944. (doi:10.1080/02699931.2014.957656)
44. 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. *Behav. Res. Methods* **40**, 531–539. (doi:10.3758/BRM.40.2.531)
45. Maurage P, Joassin F, Philippot P, Campanella S. 2007 A validated battery of vocal emotional expressions. *Neuropsychol. Trends* **2**, 63–74. (doi:10.7358/neur-2007-002-maur)
46. Lima CF, Castro SL, Scott SK. 2013 When voices get emotional: a corpus of nonverbal vocalizations for research on emotion processing. *Behav. Res. Methods* **45**, 1234–1245. (doi:10.3758/s13428-013-0324-3)
47. Amorim M, Anikin A, Mendes AJ, Lima CF, Kotz SA, Pinheiro AP. 2021 Changes in vocal emotion recognition across the life span. *Emotion* **21**, 315–325. (doi:10.1037/emo0000692)
48. Bürkner PC. 2018 Advanced Bayesian multilevel modeling with the R package brms. *R J.* **10**, 395–411. (doi:10.32614/RJ-2018-017)
49. Anikin A. 2019 Soundgen: an open-source tool for synthesizing nonverbal vocalizations. *Behav. Res. Methods* **51**, 778–792. (doi:10.3758/s13428-018-1095-7)
50. Berry DS. 1990 Vocal attractiveness and vocal babyishness: effects on stranger, self, and friend impressions. *J. Nonverbal Behav.* **14**, 141–153. (doi:10.1007/BF00996223)
51. Hughes SM, Dispenza F, Gallup GG. 2004 Ratings of voice attractiveness predict sexual behavior and body configuration. *Evol. Hum. Behav.* **25**, 295–304. (doi:10.1016/j.evolhumbehav.2004.06.001)
52. Montepare JM, Zebrowitz-McArthur L. 1987 Perceptions of adults with childlike voices in two cultures. *J. Exp. Soc. Psychol.* **23**, 331–349. (doi:10.1016/0022-1031(87)90045-X)
53. Murphy NA, Lehrfeld JM, Isaacowitz DM. 2010 Recognition of posed and spontaneous dynamic smiles in young and older adults. *Psychol. Aging* **25**, 811–821. (doi:10.1037/a0019888)
54. Oosterhof NN, Todorov A. 2009 Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion* **9**, 128–133. (doi:10.1037/a0014520)
55. Cuddy AJC, Fiske ST, Glick P. 2008 Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. *Adv. Exp. Soc. Psychol.* **40**, 61–149. (doi:10.1016/S0065-2601(07)00002-0)
56. Sutherland CAM, Rhodes G, Burton NS, Young AW. 2020 Do facial first impressions reflect a shared social reality? *Br. J. Psychol.* **111**, 215–232. (doi:10.1111/bjop.12390)
57. Vingerhoets AJJM, Bylsma LM. 2015 The riddle of human emotional crying: a challenge for emotion researchers. *Emot. Rev.* **8**, 207–217. (doi:10.1177/1754073915586226)
58. Zeifman DM. 2004 Acoustic features of infant crying related to intended caregiving intervention. *Infant Child Dev.* **13**, 111–122. (doi:10.1002/icd.344)
59. Zeifman DM. 2001 An ethological analysis of human infant crying: answering Tinbergen's four questions. *Dev. Psychobiol.* **39**, 265–285. (doi:10.1002/dev.1005)
60. Provine RR. 2004 Laughing, tickling, and the evolution of speech and self. *Curr. Dir. Psychol. Sci.* **13**, 215–218. (doi:10.1111/j.0963-7214.2004.00311.x)
61. Warren JE, Sauter DA, Eisner F, Wiland J, Dresner MA, Wise RJS, Rosen S, Scott SK. 2006 Positive emotions preferentially engage an auditory-motor 'mirror' system. *J. Neurosci.* **26**, 13 067–13 075. (doi:10.1523/JNEUROSCI.3907-06.2006)
62. Ponsot E, Burred JJ, Belin P, Aucouturier JJ. 2018 Cracking the social code of speech prosody using reverse correlation. *Proc. Natl Acad. Sci. USA* **115**, 3972–3977. (doi:10.1073/pnas.1716090115)
63. O'Connor JJM, Barclay P. 2018 High voice pitch mitigates the aversiveness of antisocial cues in men's speech. *Br. J. Psychol.* **109**, 812–829. (doi:10.1111/bjop.12310)
64. Torre I, White L, Goslin J. 2016 Behavioural mediation of prosodic cues to implicit judgements of trustworthiness. In *Proc. of the Int. Conf. on Speech Prosody*, pp. 816–820. Boston, MA: International Speech Communication Association. (doi:10.21437/speechprosody.2016-167)
65. Young AW, Frühholz S, Schweinberger SR. 2020 Face and voice perception: understanding commonalities and differences. *Trends Cogn. Sci.* **24**, 398–410. (doi:10.1016/j.tics.2020.02.001)