Please cite as:

Wołoszyn, K., Hohol, M., Kuniecki, M., & Winkielman, P. (2024). Facing emotional vocalizations and instrumental sounds: Sighted and blind individuals spontaneously and selectively activate facial muscles in response to emotional stimuli. *Emotion*, *25*(4), 827–840. https://doi.org/10.1037/emo0001475

Facing emotional vocalizations and instrumental sounds: Sighted and blind individuals spontaneously and selectively activate facial muscles in response to emotional stimuli

Kinga Wołoszyn[1*], Mateusz Hohol[2], Michał Kuniecki[3], Piotr Winkielman[4]

[1] Institute of Psychology, Jesuit University Ignatianum in Krakow, Poland

[2] Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Krakow, Poland

[3] Institute of Psychology, Jagiellonian University, Krakow, Poland

[4] Department of Psychology, University of California San Diego, La Jolla, USA

[*] Corresponding author: kinga.b.woloszyn@gmail.com

ORCIDs:
KW: 0000-0003-2059-9914
MH: 0000-0003-0422-5488
MK: 0000-0001-6910-8667
PW: 0000-0003-2330-1802

Acknowledgments:
This study was supported by the grant “Preludium,” 2017/25/N/HS6/01052 (PI: Kinga Wołoszyn, Supervisor: Piotr Winkielman) funded by the National Science Centre, Poland. MH was supported by the Priority Research Area “Society of the Future” of the Strategic Programme ‘Excellence Initiative’ at Jagiellonian University. We would like to thank Anna Jankowska for helping in the recruitment of participants.

Abstract:

Facial mimicry of visually observed emotional facial actions is a robust phenomenon. Here, we examined whether such facial mimicry extends to auditory emotional stimuli. We also examined if participants' facial responses differ to sounds that are more strongly associated with congruent facial movements, such as vocal emotional expressions (e.g., laughter, screams), or less associated with movements, such as non-vocal emotional sounds (e.g., happy, scary instrumental sounds). Furthermore, to assess whether facial mimicry of sounds reflects visual-motor or auditory-motor associations, we compared individuals that vary on lifetime visual experience (sighted vs. blind). To measure spontaneous facial responding, we used facial electromyography to record the activity of the *corrugator supercilii* (frowning) and the *zygomaticus major* (smiling) muscles. During measurement, participants freely listened to the two types of emotional sounds. Both types of sounds were rated similarly on valence and arousal. Notably, only vocal, but not instrumental, sounds elicited robust congruent and selective facial responses. The facial responses were observed in both sighted and blind participants. However, the muscles' responses of blind participants showed less differentiation between emotion categories of human vocalizations. Furthermore, the groups differed in the shape of the time courses of the *zygomatic* activity to human vocalizations. Overall, the study shows that emotion-congruent facial responses occur to non-visual stimuli and are more robust to human vocalizations than instrumental sounds. Furthermore, the amount of life-time visual experience matters little for the occurrence of cross-channel facial mimicry, but it shapes response differentiation.

Keywords: emotional sounds; fEMG; embodiment; mimicry; blind

1. Introduction

Human facial displays can reflect and communicate our affective states and intentions and thus play an important role in everyday social interactions. However, despite years of research, some fundamental questions about these displays are still unresolved. These questions relate to the nature of stimuli triggering facial displays, specificity and timing of underlying muscle movements, as well as the role of the displayer's visual experience in shaping those movements. The current research examined these questions by comparing facial responses, measured using electromyography (EMG), to non-visual emotional stimuli: sounds. We used two types of sounds conveying emotions: human vocalizations and instrumental sounds. The goal of this part of our investigation was to test if spontaneous and emotion-selective facial reactions occur to non-visual emotion stimuli and if the shape of these reactions depends on whether a stimulus is generated by emotion-related human facial activity (vocalization) or is generated via other means (instrumental sounds). The second and related goal of the study was to address the role of the displayers' visual experience in facial reactivity to both kinds of affective sounds, through comparing selectivity and temporal dynamics of reactions in sighted and blind individuals. By addressing these two goals, the current research informs the theoretical debates about the nature and mechanism of spontaneous facial responding, as discussed next.

1.1. The nature and nurture of facial expressions

There is a long-standing debate on facial expressions. An influential evolutionary perspective proposes that facial expressions selectively reflect or communicate discrete emotions, correlate with distinct subjective states, and have partially innate, fairly fixed patterns that are produced as well as recognized across cultures (Cowen et al., 2021; Darwin, 2009; Ekman, 1993; Matsumoto et al., 2008). Other views emphasize the role of local ecology, social learning, and the specific cultural context in shaping both the production and perception of facial expressions

(Barrett et al., 2019; Russell & Fernandez-Dols, 1997), see also (Fridlund, 1994; Keltner, 1995; Tiedens et al., 2000; Weisbuch & Adams Jr., 2012). While there is a lively debate about facial displays' origin and primary function, it is widely agreed that facial expressions play an important role in social life as they reflect and communicate affective states and intentions to others.

1.2. Spontaneous facial expressions and emotional mimicry

One class of facial displays that is particularly interesting to researchers are spontaneous facial expressions, as opposed to posed, voluntary, expressions. The assumption is that spontaneous expressions are more likely to reflect quicker, lower-level, and less-strategic processes. These spontaneous expressions can occur to a variety of stimuli, including others' faces, and, when their form matches the eliciting stimulus, they are often termed "mimicry" (Dimberg, 1982; Hess & Fischer, 2013). We will later address in more detail the question of the difference between facial actions constituting *motor mimicry*, reflecting a mere motor match, *social mimicry*, reflecting an other-oriented response, and *emotional response*, reflecting an underlying emotional state elicited by the stimulus. For now, however, note that so-called *emotional mimicry* is an example of the broader phenomenon. Mimicry, in general, is a tendency to spontaneously replicate the observable behavior of others (Arnold & Winkielman, 2020; Chartrand & Lakin, 2013). As such, mimicry can refer to multiple kinds of actions, including non-emotional facial actions such as mouth opening or yawning (Palagi et al., 2020), or even non-facial neutral actions such as foot tapping, leg crossing (Ashton-James et al., 2007; Chartrand & Bargh, 1999), and body postures (Bavelas et al., 1986; Bernieri & Rosenthal, 1991). However, for our purposes here, we will focus on the literature on facial responses to emotional stimuli.

Spontaneous facial mimicry to emotional faces has been shown in a number of studies using facial electromyography (fEMG; Dimberg et al., 2000; Kret et al., 2013; Tamietto et al., 2009) and methods such as the Facial Action Coding System (FACS; Ekman et al., 2002; Ekman & Friesen, 1978) or the Maximally Discriminative Facial Movement Coding System (MAX; Izard, 1979). The mimicry of facial muscle patterns occurs quickly (Dimberg, 1982; Lundqvist & Dimberg, 1995) and the perceiver does not need to be aware of one's own facial changes or even mimicked expressions (Bornemann et al., 2012; Dimberg et al., 2000; Moody et al., 2007; Tamietto et al., 2009). At the same time, Murata and colleagues (2016) demonstrated that this largely automatic process may depend on whether the individual is oriented toward inferring the emotional states of the person whose expression is being observed.

Research on facial mimicry – with its focus on a matching response – has been mostly focused on responses to others' facial expressions. However, emotions can also be conveyed and elicited by stimuli in other modalities. Emotions expressed with sounds are well recognized. This is the case with voice prosody (Banse & Scherer, 1996; Leinonen et al., 1997; Scherer, 1981; Scherer et al., 1991), and also with non-verbal affect bursts (Belin et al., 2008; Hawk et al., 2009; Schroder, 2003; see Kamiloğlu et al., 2021 for the comparison of prosody and non-verbal vocalizations). Some research using self-reports found that in everyday communication, vocal expressions sometimes convey emotions better than facial expressions (Planalp, 1996). In the more recent study by Hawk and colleagues (2009), emotions expressed with simple affect vocalizations were recognized as accurately as facial expressions. Furthermore, emotions such as anger, contempt, disgust, fear, sadness, and surprise were recognized even more accurately from non-verbal vocalizations than from facial displays. Moreover, in the study by Kamiloğlu and collaborators (2021) participants recognized various positive emotions more accurately when expressed via non-verbal vocalizations than prosody.

Several studies have demonstrated that emotional sounds evoke relevant facial displays, analogously to facial expressions. In the study by Dimberg (1990) that focused on the dimension of valence, simple unpleasant tones elicited greater activity of the *corrugator supercilii*, the muscle associated with negative valence. Similarly, Bradley and Lang (Bradley & Lang, 2000) used emotion-evoking sounds and observed a standard pattern of facial response, namely, increased activity of the *zygomaticus major* while listening to pleasant sounds and greater activity of the *corrugator supercilii* to unpleasant sounds.

Addressing this issue in an innovative way, Arias and colleagues (2018, 2021) artificially transformed the neutral sentences to include previously identified smile-specific acoustic cues in prosody (stress and intonation). Their study revealed the activation of the *zygomaticus major* and the relaxation of the *corrugator supercilii* to transformed sentences, even when participants did not consciously recognize those sentences as spoken with a smile. This study highlights that the oro-facial configuration (shape of the human mouth and face) of the sound producer can shape the facial response of the listener, presumably via mechanisms linking perception and action. That is, even though musical and human expressions share certain acoustic properties, which differentiate emotional and neutral sounds, the study by Arias et al. (2018) showed that what triggers smiling in response to human vocal expressions is a cue resulting from the specific human oro-facial activity at production.

Note that the just discussed studies primarily focused on the broad dimension of positive-negative valence. This limitation does not apply to experiments which focused on specific emotion categories by Hawk and colleagues (2012). To examine the participants' facial expression in more detail, the authors used the FACS (Ekman et al., 2002; Ekman & Friesen, 1978). In their study, the authors used non-verbal expressions of anger, disgust, sadness, joy, as well as neutral sounds, represented by grunts, retching, laughter, sobs, and cries, among others. The sounds triggered congruent facial expressions in listeners, including muscle movements specific not only to the valence of the stimulus, but also to a given emotion category (e.g., participants wrinkled their noses to disgust-expressing stimuli more often than to the other of emotion categories). Following up on earlier demonstrations of selective facial movements to emotional sounds, Vilaverde and colleagues (2024) recently showed that interfering with such facial responses through restricting facial movements hinders a process of differentiating authentic from posed vocal expressions, i.e., laughs and cries.

1.3. Emotion response vs. emotion mimicry

One broad challenge in mimicry research is whether the observed facial response reflects a purely motor mimicry (matching motor response), a more socially-oriented mimicry (motor response directed at others), or an elicited emotional state (but see Shaham & Aviezer, 2022). This challenge is particularly relevant in the context of sounds because there is seemingly no match between the physical form of the auditory stimulus and the facial response (e.g., compared to a smiling response to a picture of a smile). However, Hawk and Fischer (2016) argue that mimicry goes beyond a single modality or a simple match in stimulus-response format. This is because people usually respond to the *meaning*, rather than a mere form, of a stimulus. In fact, recent research on facial mimicry to facial stimuli has shown that participants can “mimic” invisible (due to occlusion) parts of other’s faces. For example, in social contexts, participants spontaneously activate their own cheek muscles even if they only see the other person’s smiling eyes (because the other person’s cheeks are occluded). Interestingly, in a non-social context which deemphasized the role of observers’ focus on expression meaning, mimicry was limited to specifically observed muscles (Davis et al., 2017, 2022). These findings are consistent with a general framework emphasizing that in social contexts mimicry often goes beyond a simple motor match (Hess & Fischer, 2014). Consistent with this social view of mimicry, people are more likely to mimic emotions that promote affiliation (e.g., happiness or sadness) than antagonizing emotions such as anger (Bourgeois & Hess, 2008; Fischer et al., 2012; Hinsz & Tomhave, 1991; Olszanowski et al., 2020). This tendency for stronger response to positive expressions has also been shown in the mimicry of vocal expressions of emotions (Hawk et al., 2012). This is important for two reasons. First, this work highlights that even spontaneous facial reactions reflect not only the motor configuration of the stimulus, the valence of the stimulus, but also who is generating the stimulus and the possible social function of the facial reaction (e.g., amplification of smiling to positive stimuli). Second, and related, this work highlights that responses to positive stimuli may be unique in the social context as positive stimuli may be particularly worth "imitating" (Hess & Fischer, 2013). Both of these points become relevant when considering potential differences between human-generated vocalizations and emotional sounds generated by instruments, and also to the later discussion between facial responses of people with different amounts of visual experience, such as the sighted and the blind, which we compare in the current research.

People convey arousal, valence, and discrete emotions not only through their bodies but also through cultural artifacts, including music (Juslin, 2013; Juslin & Laukka, 2003; Murray et al., 1996; Scherer, 2004). The review by Juslin and Laukka (2003), which included nearly 150 studies, revealed that vocal and musical expressions are categorized in terms of discrete emotions as accurately as human vocalizations. This high recognition rate presumably stems from shared acoustic characteristics of vocal/instrumental sounds. Moreover, some studies suggest that emotions conveyed by the timbre of an isolated musical instrument in some way resemble speech prosody (Liu et al., 2018), connecting these studies to the recent work on the effects of prosody as mentioned earlier (Arias et al., 2018). Single studies explored the facial responses to positively and negatively valenced musical pieces and found increased activity of the *zygomaticus major* during listening to happy instrumental excerpts (Kallinen, 2004; Lundqvist et al., 2009). This result is generally in line with the results of the studies using social and non-social emotional sounds (e.g., Bradley & Lang, 2000 using erotica, bombs, and screams), though the exact comparison is missing.

Although instrumental sounds can convey emotions, evoke facial responses, and can be a part of social context, they are clearly distinct. One difference is that, unlike facial or vocal expressions of emotion, instrumental sounds do not provide direct information about the affective state of an interaction partner, whose emotion we need to recognize, understand, then accurately respond to in the case of real emotion displays. To give a simple example, hearing sad instrumental music may be clearly understood as sad, or even make the perceiver sad. However, it does not imply the musician is sad, and thus does not elicit similar motivation for mimicry. This is different from contact with a sad face or a crying voice (Hess & Fischer, 2013). Therefore, facial responses to emotional music are less likely to be driven to the same extent by social motivation for mimicry as emotion to human sounds. Another important difference between instrumental sounds and human vocalizations is the potential role of associations between production and perception – an issue we discuss next.

1.4. Mimicry and the amount of visual experience

One dominant explanation for mimicry of visual emotional stimuli – the link between seeing a facial expression (e.g., noticing a smile) and producing it oneself (e.g., smiling back) is visual-motor matching. In essence, a visual input elicits congruent motor output (Prinz, 1990). But what is the basis for the link between hearing an emotion vocalization and responding with congruent facial display? After all, the auditory input and motor output do not match in form. Does this link require visual experience? According to one view, the visual experience might indeed be, at least at some point, needed for establishing the association between emotion vocalizations and congruent facial responding. Along with hearing emotional vocalizations, we also often see a person's face producing them. This could lead to an establishment of a link between visual perception and motor action link, which would support automatic imitation, even across modalities (Catmur et al., 2009). From this perspective, a strong link between the affective sounds and facial responses would require visual experience. From another perspective, the association between affect vocalizations and facial responding might be independent from any visual experience and be more linked to sound production. After all, people engage appropriate facial muscles while producing emotional vocalization. As Hawk and colleagues (2012) demonstrated, facial muscle activity present while producing emotion vocalizations is correlated with the one occurring during the perception of sounds expressing those emotions. These findings suggest that the auditory-motor association might be created via our own auditory-proprioceptive experience. Viswanathan and colleagues (2024) described the scenario of the ontogenetic development of facial imitation, which is consistent with this proposition. Infants vocalize frequently, also when there is no one around. While vocalizing, they learn to associate the spectro-temporal features of the sound resulting from making a particular oro-facial gesture with their own motor facial action and proprioceptive signals. Visual modality comes second when, during dyadic interactions with a caregiver, a child learns to associate visually perceived facial expressions with a motor movement triggered by the already-known sound. Therefore, facial responses to emotional vocalizations do not require any visual experience.

According to the studies by Arias and colleagues (2018, 2021), which were focused on the expressions of happiness, happy facial displays, particularly the contraction of the *zygomaticus major*, cause the specific formant shift in the sound. This auditory cue, thus, provides direct information about the oro-facial configuration of the producer, which triggers the corresponding muscular activity in the perceiver. The claim of Arias and colleagues has been supported by their recent study (Arias et al., 2021), in which blind participants (including congenitally blind) responded with the contraction of the *zygomaticus* to neutral sentences transformed to include smiling-specific formant shift. Reflecting the potential importance of visual experience in mimicry, our study aimed to shed more light on this issue and compared sighted and blind participants in terms of their facial responses to affect vocalizations, including a wider range of emotion categories, and offering a direct comparison between vocal and instrumental sounds. In addition, like Arias and colleagues, we compared people who vary on the amount of visual experience.

1.5. Facial displays in the blind

Valente and colleagues (2018) reviewed 21 studies published between 1935 and 2015, which focused on spontaneous and voluntarily posed facial expressions among blind people. Concerning spontaneous expressions, correlational and experimental studies on both children and adults suggest no substantial difference in spontaneous facial responses produced by blind or sighted individuals. However, most studies included in the review by Valente et al. (2018) concerned complex social situations that evoked emotions, like winning a gold medal (Matsumoto & Willingham, 2009) or interpreting old sayings during structured interviews (Rinn, 1991). Furthermore, the expressions classified as spontaneous in those studies might also be considered as influenced by control mechanisms or “partially posed” in a sense that they were displayed while participants were being watched by other people. In fact, according to

Matsumoto and colleagues (Matsumoto et al., 2009; Matsumoto & Wilson, 2023), when a person is aware of being looked at, the immediate, initial expressions are followed by the culturally shaped displays. Therefore, these studies might still focus on the socially shaped responses, while overlooking the very early, initial, natural expressions that are generated in a more anonymous setting. Therefore, to observe genuinely spontaneous expressions, in a setting that minimizes the sense of being watched by others, researchers have used measures that enable unobtrusive detection of subtle, quick changes in facial movements, like fEMG. In this vein, one study by Arias and colleagues (2021) showed that blind individuals responded with the *zygomatic* contraction to neutral sentences "read with a smile," as sighted participants. It remains unknown, however, whether blind people present emotional mimicry to a wider range of categories of auditory expressions in terms of emotion categories and mode of production (vocal vs. instrumental).

1.6. Objectives of the present study

In the present study, we investigated facial responses to auditory emotion stimuli. Our first goal was to test whether participants would show congruent facial muscle responses to brief auditory stimuli expressing fear, sadness, happiness, and neutral sounds and whether facial responses would differ between human affect vocalizations and instrumental sounds expressing the same emotion categories. We predicted that emotional sounds would evoke facial responses congruent with the sound category and that the responses will be stronger in the case of vocalizations compared to instrumental sounds.

Secondly, to shed light on the issue of visual experience in the establishment of the association between auditory emotion stimuli and congruent facial responding, we examined whether those facial responses are similar between sighted and blind individuals. The extant theoretical considerations allow for two different predictions regarding the importance of visual experience. If a link between auditory perception and one's own production of facial response is key, blind individuals's reaction should be similar to sighted people, especially for the vocalization. After all, blind people use the same facial muscles to generate vocalizations.

If a link between visual perception and facial response is key, the sighted group should respond more robustly than the blind. This is grounded in the idea that sighted individuals have more experience between visual perception of the facial expression and sounds (e.g., hearing a laugh is associated with seeing a smile). Finally, because musical expressions do not have a close link between the sounds and the face-related visual, motor, and proprioceptive cues, we should not observe any differences between sighted and blind individuals in terms of facial responding to this sound category.

We presented a set of stimuli expressing fear, sadness, happiness, and neutral sounds to sighted and blind participants. The stimuli represented two types: human-made vocal emotion expressions (e.g., laughter, scream) and non-vocal emotional sounds (brief instrumental sounds). Participants were freely listening to the sounds, with no additional task, to reflect more natural facial responses. We measured the activity of the *corrugator supercilii* (frowning) and the *zygomaticus major* (smiling) muscles with facial electromyography.

2. Method

2.1. Participants

Thirty-five sighted individuals participated in the study. Two were excluded from the analyses due to technical problems during data collection, and one because of the excessive number of artifacts in the EMG signal. The remaining group of 32 participants had an average age of 31.34

years (*SD* = 5.82) and included 17 women. The second group consisted initially of 20 participants with total or severe vision loss without comorbid psychiatric or neurological disorders. One person was excluded due to a technical problem with the electrode attachment, and another because of too many artifacts in the electromyographic data. Out of 12 participants who agreed to provide the data, 6 were congenitally blind, 3 participants had severe vision loss since birth, one lost their sight in childhood, one in adolescence, and one as an adult. No participants reported cortical damage – vision loss was due to the eye or optic nerve damage. Single participants reported being able to see the light, big objects, or bright colors. None of them was able to recognize faces. The remaining group consisted of 18 participants (M_{age} = 32.5; SD_{age} = 6.39; 7 female). The size of this group was limited by the recruitment capacity of people from this specific population. Participants from the sighted group were recruited through advertisements posted on the Internet. The second group consisted of persons associated with the Seventh Sense Foundation and Katarynka Foundation through advertisements posted on social media, and newsletters. All participants were Polish. They received a financial reward and provided informed consent (as explained in the Procedure). The design of the study was approved by the Ethics Committee for Experimental Research at the Institute of Psychology, Jagiellonian University in Krakow (decision number KE/04/062018).

2.2. Experimental materials

We used 80 short sounds that were non-verbal expressions of three basic emotions: fear, sadness, and happiness, and 20 neutral sounds. Half of the stimuli (40) were selected from the Montreal Affective Voices (MAV; Belin et al., 2008), consisting of brief human-made affect vocalizations (e.g., laughing, crying). Half of them were male and half female. The sounds lasted on average 1.33 seconds (*SD* = 0.81). The other 40 sounds came from the Musical Emotional Bursts database, which was created as a musical analog of the MAV (MEB; Paquette et al., 2013). Instrumental sounds were brief melodies played on either the violin or the clarinet. They lasted, on average, *M* = 1.6 seconds (*SD* = 0.64). According to the validation made by the authors of the databases (Belin et al., 2008; Paquette et al., 2013), the sounds we selected for our study were characterized by high recognizability, i.e., emotions expressed by those sounds were recognized in over 91% (happy sounds had the highest recognizability at 98.25%, fear was the lowest at 87.25%).

2.3. Procedure

Upon arrival at the laboratory, the experimenter presented the participants with the general purpose of the study and their task, and he/she explained the procedures, including the muscle sites' preparation and placement of the electrodes. All subjects were informed that they would be video recorded during the experiment. Then, participants gave their informed consent. Sighted participants did it in writing. Blind participants were read the form and gave their consent verbally. Their declarations were recorded.

Before the electrodes' attachment, the experimenter removed any makeup, applied the abrasive paste to the electrode sites, and cleaned them with salicylic alcohol. Using bipolar placement, the Ag/AgCl electrodes were placed with saline-based electrode gel over the *corrugator supercilli* and the *zygomaticus major* regions on the left side of the face, according to the guidelines provided by Fridlund and Cacioppo (1986). A ground electrode was placed below the hairline in the middle of the forehead. The study consisted of two procedures. In both, the presentation of the stimuli was controlled by the PsychoPy software (Peirce, 2009).

The first task was to listen freely to 80 sounds. Each sound was followed by a period of silence lasting randomly from 10 to 15 seconds. The sounds were presented in a random order through the speakers placed at the desk on both sides of the participant. The participants adjusted the sound volume to a comfortable value before the first task, and it stayed the same until the end of the study. Overall, the task lasted approx. 25 minutes. Participants were filmed during this part of the study to enable later offline artifact rejection.

The second procedure was evaluating the sounds on the valence and arousal scales using the Self-Assessment Manikin (SAM; Bradley & Lang, 2000). During this procedure, the EMG data were not recorded. Participants rated each sound on the negative to positive dimension (*To what extent were the emotions evoked by the sound negative or positive?*) and on the calming to arousing dimension (How calming or arousing *were the emotions evoked by the sound?*) using a 9-point Likert-type scale ranging from the most negative (1) to the most positive (9) and from the most calming (1) to the most exciting (9). Sighted participants answered by clicking on the SAM rating scales after each sound with a computer mouse. The emotional judgment of participants with vision loss was rated using a tactile version of the SAM created by Iturregui and Méndez-Ulrich (T-SAM; 2020). The T-SAM is a simplified version of the SAM prepared using the relief printing technique. Before the evaluation procedure, the experimenter helped blind participants become acquainted with the scale by touch, giving them as much time as needed to get familiar with the tool. After each sound, the experimenter asked the participants about their valence and arousal ratings, reminding them what each end of the scale meant. Participants gave their answers orally, and the experimenter recorded the values. After the experiment, the experimenter debriefed all participants as to the purpose of the study.

2.4. Data acquisition, preprocessing, and software

The activity from the *corrugator supercilii* and *zygomaticus major* sites was recorded using a wireless BioSignalsPlux toolkit (PLUX Wireless Biosignals) and a custom LabView program.

The signal was digitized with 16-bit resolution, sampled at 2000 Hz, and filtered with a 25-500 Hz anti-aliasing bandpass filter.

The preprocessing of the data was conducted using BrainVision Analyzer (BrainVision Analyzer, Version 2.2.0, Brain Products GmbH, Gilching, Germany). First, the raw EMG data were visually inspected, along with a video recording, to mark artifacts. All fragments including artifacts (e.g., stemming from coughing, yawning), were excluded from the analysis. The data of two participants was removed due to an excessive number of segments, including artifacts (25 segments). Then, the data was filtered with a 10–500 Hz bandpass filter, rectified, and smoothed using a moving average transform with a 125.5 ms window size. In the next step, the data were divided into 6 s segments—from 0.5 s before to 5.5 s after stimulus onset—and standardized. Then, the baseline correction was applied—the average from the 0.5 s period before the stimulus onset was subtracted from each subsequent data point from a given trial. All the data was down-sampled to 2 Hz (giving 12 data points per trial) and exported for further statistical analysis.

To analyze temporal dynamics of facial muscles reactions (repeated-measures ANOVAs) we used SPSS 26 (IBM Corp. Armonk, NY). To analyze facial activity indexes with a linear mixed model (LMM), we used R-based software (R Core Team, 2023), i.e., Jamovi 2.2 (The Jamovi project, 2023) with a module *GAMLj* (Gallucci, 2019), and RStudio (RStudio Team, 2023) with packages *lmerTest* (Kuznetsova et al., 2017) and *effectsize* (Ben-Shachar et al., 2020) to calculate effect sizes.

2.5. Transparency and openness

We report all data exclusions, all manipulations, and all measures in the study. The data collected during the current study are available at the Open Science Framework (https://osf.io/tpdmz/). This study's design and its analysis were not pre-registered.

3. Results

3.1. Facial activity indexes

As emphasized by Hess and colleagues, facial mimicry is the spontaneous adaptation of a perceived configuration of muscle movements and not the response of a single muscle (Hess & Blairy, 2001; Hess et al., 2017; Olszanowski et al., 2020). Thus, following this approach, we used a facial activity index, which considers the activity of both the *corrugator* and the *zygomaticus*. The index is calculated by subtracting the standardized *corrugator* activity from the standardized activity of the *zygomaticus*. These muscles have been shown to reflect the responses most reliably to positive and negative stimuli. The higher the index, the stronger the activity of the *zygomaticus* relative to the *corrugator*. Then, based on facial indexes, we used a linear mixed model with restricted maximum likelihood estimation and included emotion (fear, sad, happy, neutral), sound type (vocal, instrumental), and group (sighted, blind) as fixed factors and intercept fit across participants and stimuli as random factors (index ~ group + emotion + sound_type + 1|participant + 1|stimulus). Bonferroni correction was used for post-hoc comparisons. As expected**,** the analysis yielded a significant effect of emotion ($F(3, 80.6) = 11.03$; $p < .001$; $\eta_p^2 = 0.29$). The index for happy sounds ($M = 0.28$) was significantly higher than the indexes for all other emotion categories, that is, sadness ($M = 0.05$; $p < .001$), fear ($M = 0.04$; $p < .001$), and neutral ($M = 0.07$; $p < .001$). There was also a significant effect of sound type ($F(1, 80.6) = 13.22$; $p < .001$; $\eta_p^2 = 0.14$) and the interaction of sound type and emotion ($F(3, 80.6) = 3.52$; $p = .019$; $\eta_p^2 = 0.12$). Human vocalizations were linked to higher facial indexes than instrumental sounds ($M = 0.17$ vs $M = 0.06$). As indicated by the results of the post-hoc tests, this difference was significant only for happy sounds ($p = .002$). Notably, the differences in facial indexes among emotion categories, as those observed for the main effect of emotion, were significant only in the case of human vocalizations (all $p < .01$).

As for the between-group differences, the main effect of the group was insignificant ($F(1, 46.7) = 1.65$; $p = .206$; $\eta_p^2 = 0.03$). However, the analysis yielded a significant interaction between group and emotion ($F(3, 3431.4) = 3.41$; $p = .017$; $\eta_p^2 = 0.003$). The difference between the groups concerned the facial responses to sad sounds. In the sighted, the value of the index for sad sounds was significantly lower than its value for happy sounds ($p < .001$), and in the blind, there was no such difference ($p = 1.00$). Further, there was no significant interaction between the group and sound type ($F(1, 3430.3) = 1.88$; $p = .170$; $\eta_p^2 < 0.001$), as well as group, sound type, and emotion category ($F(3, 3430.7) = 0.12$; $p = .951$; $\eta_p^2 < 0.001$). The facial activity indexes for four emotion categories divided by two sound types are shown in Fig. 1.

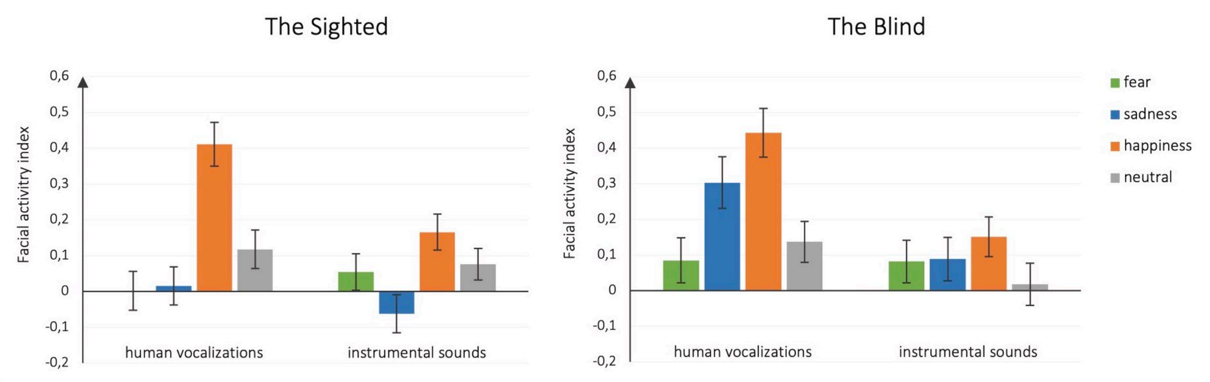


Fig 1. The facial activity indexes for four emotion categories and two sound types for sighted (left) and blind (right) participants separately. The facial indexes are calculated by subtracting the z-scored mean activity of the *corrugator supercilii* from the z-scored activity of the *zygomaticus major*. The higher the index, the greater the *zygomaticus* activity relative to the *corrugator* activity. Error bars represent standard errors.

3.2. Temporal dynamics of facial muscles reactions

We compared the temporal dynamics of both muscles' activity during the 5.5s period which began at the sound presentation. First, for each group and each muscle separately, we conducted repeated-measures ANOVAs with factors of emotion (fear, sad, happy, neutral), sound type (vocal, instrumental), and time (12 time points). In both groups, there was a significant effect of emotion on the time courses of the activity of both muscles, the *corrugator* (sighted: $F(33, 1023) = 2.89$; $p < .001$; $\eta_p^2 = 0.09$; blind: $F(33, 561) = 1.49$; $p = .040$; $\eta_p^2 = 0.08$) and the *zygomaticus* (sighted: $F(33, 1023) = 2.76$; $p < .001$; $\eta_p^2 = 0.08$; blind: $F(33, 561) = 1.78$; $p < .005$; $\eta_p^2 = 0.10$). In the case of the *corrugator*, in both groups, there was no significant interaction between time and sound type (sighted: $p = .292$; blind: $p = .334$) as well as time, emotion, and sound type (sighted: $p = .081$; blind: $p = .717$). In the case of the *zygomaticus*, only in the blind, the interaction of time and sound type ($F(11, 187) = 4.08$; $p = .002$; $\eta_p^2 = 0.19$), as well as time, sound type, and emotion were significant ($F(33, 561) = 1.74$; $p < .007$; $\eta_p^2 = 0.09$).

To compare the groups, we conducted mixed ANOVAs with between-subject group factor (2) and within-subject factors of time (12), emotion (4), and sound type (2). In the case of the violation of the sphericity assumption, the Huynh-Feldt correction was applied. The analysis showed a significant group x time x emotion interaction in terms of the activity of the *zygomaticus* ($F(33, 1584) = 2.01$; $p = .006$; $\eta_p^2 = .04$). The interaction of group, time, and sound type ($p = .122$), as well as group, time, sound type, and emotion ($p = .174$) in terms of the activity of the *zygomaticus*, were insignificant. There were no differences between the groups regarding the time courses of the *corrugator* activity. The time courses of the *zygomaticus* and the *corrugator* responses are presented in Figures 2 and 3, respectively.

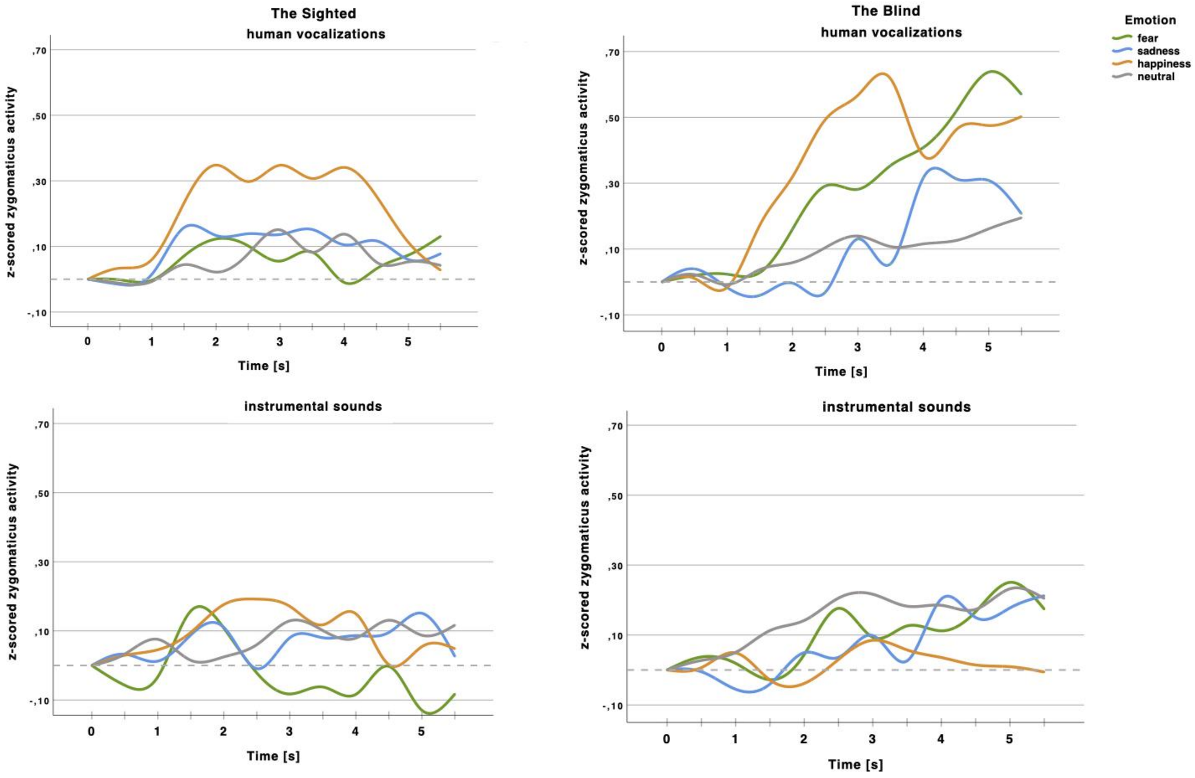


Fig 2. *Z*-scored zygomaticus activity in the time window from stimulus presentation to 5.5 seconds later. In the separate graphs, we present the responses to vocal vs. instrumental affect sounds for sighted (left side) and blind (right side) participants separately. See Supplementary Materials, Section 3 for the plots including CI.

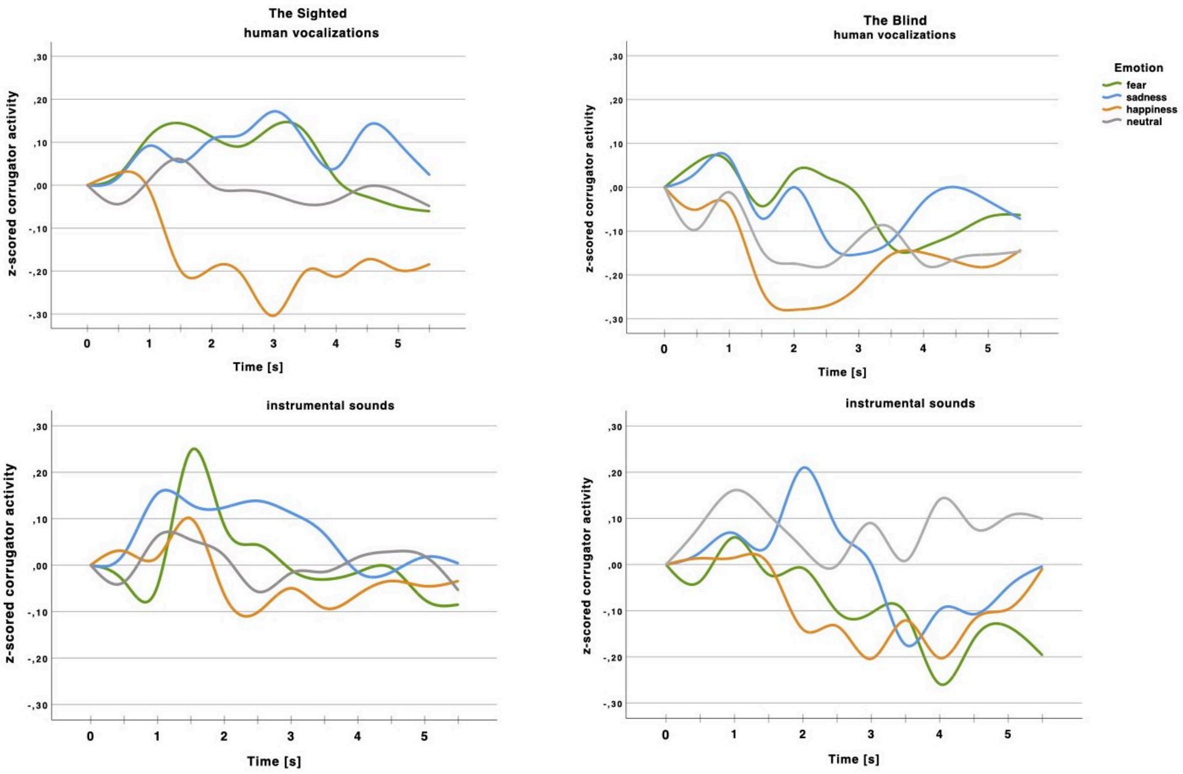


Fig 3. *Z*-scored corrugator activity in the time window from stimulus presentation to 5.5 seconds later. In the separate graphs, we present the responses to vocal vs. instrumental affect sounds for sighted (left side) and blind (right side) participants separately. See Supplementary Materials, Section 3 for the plots including CI.

3.3. Stimuli evaluation

In terms of valence rating, several patterns are worth pointing out. First, overall the vocal sounds were rated similarly to instrumental sounds. This is worth emphasizing because of the marked differences in EMG responses between the two broad classes of affective sounds. In terms of detailed comparisons, we found several local differences both between the groups and between vocal and instrumental emotional expressions within each group. First, sighted participants rated fearful vocal expressions as less negative than instrumental expressions, and blind participants rated both types of stimuli as nearly equally negative. On the other hand, for sighted participants, sad and happy sounds did not differ between sound types. In the blind, sad human vocalizations were rated as more negative than sad instrumental excerpts, and happy instrumental melodies were rated as more positive than happy vocalizations. The valence ratings of the two types of sounds in both groups are presented in Figure 4 and Table 1. All descriptive statistics and the results of further comparisons concerning the valence and arousal ratings are presented in the Supplementary Materials, Section 4.

Notably, the rating differences in self-reports are not reflected in the EMG data. Regarding the sighted, there was no difference in the facial activity index between vocal and instrumental stimuli for fear-expressing vocalizations, which were present for self-report data. In the blind, a significant difference in facial activity indexes concerned the happy sounds, but the difference was in the opposite direction to that of the valence ratings—human happy vocalizations, which were rated as less positive, evoked a greater *zygomatic* response than instrumental sounds, which were, in turn, rated higher on the valence scale. The discrepancies in the rating of sad stimuli present in the blind were not accompanied by any difference in terms of facial activity. To test the claim that the effects in terms of facial activity indexes are not due to the differences in self-reported ratings, we used a linear mixed model, in which we included those ratings. It confirmed our thesis. The results are presented in the Supplementary Materials, Section 5).

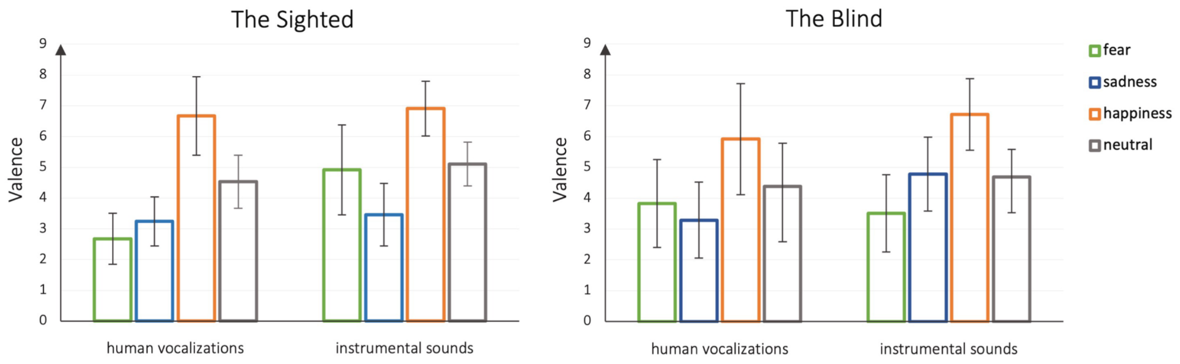


Fig 4. The valence ratings for four emotion categories and two sound types for sighted (left) and blind (right) participants separately. The scale ranges from 1 (the most negative) to 9 (the most positive).

Table 1. The comparisons of valence ratings between vocal and instrumental stimuli in sighted blind individuals.

Group		vocalizations		instrumental sounds		difference	
		Mean	*SD*	Mean	*SD*	*t*	*p*
sighted							
	Fear	2.68	0.83	4.92	1.46	-2.24	<.001
	Sad	3.24	0.8	3.46	1.02	-.22	.19
	Happy	6.67	1.28	6.91	0.89	-.24	.425
	Neutral	4.53	.86	5.11	0.71	-.59	.015
blind							
	Fear	3.83	1.42	3.51	1.25	.33	.3
	Sad	3.29	1.23	4.78	1.2	-1.48	<.001
	Happy	5.92	1.8	6.72	1.16	-.8	.015
	Neutral	4.39	1.39	4.69	0.89	-.3	.25

4. Discussion

4.1. Overview

Our study used electromyography to investigate spontaneous facial responses to auditory stimuli representing fear, sadness, happiness, and neutral emotion categories. There were two types of stimuli which were matched on self-reported valence and arousal: emotional vocalizations (e.g., laughter, scream) and non-vocal brief music excerpts (played on violin, clarinet). We expected to observe facial responses congruent with the sounds' emotion category. Furthermore, we expected more pronounced and differentiated facial responses to human vocalizations. This is because, compared to instrumental excerpts, the production of vocalizations involves selective activity of facial muscles, which may result in a stronger perception-production link. The secondary goal of the study was to explore the role of visual experience in the link between affective sounds with facial responses. Thus, we recruited sighted and blind individuals. If facial mimicry to affect vocalizations relies on visual-motor associations, sighted participants should show more pronounced effects. However, if auditory-motor associations are key, both groups should show similar facial mimicry to affect vocalizations. Furthermore, since, compared to affect vocalizations, instrumental excerpts are less associated with congruent face-related production activity, we expected less differentiation in facial responses to instrumental excerpts in both sighted and blind participants.

4.2. Facial responses to auditory emotion conveyed by human vocalizations vs. instrumental sounds

As expected, the pattern of facial activity depended on the emotion conveyed by the sound. The index of positive facial response (greater activity of the *zygomaticus major* relative to the *corrugator supercilii*) was higher in response to happy sounds than to neutral sounds and sounds expressing negative emotions, namely fear and sadness. Notably, the described effect of emotion on facial responses was robust for human affect vocalizations, but not for instrumental sounds. Importantly, this difference in facial response does not seem to result from the weaker intensity or greater ambiguity of the instrumental excerpts. First of all, on self-reports our participants evaluated human vocalization and instrumental sounds largely similarly. This finding is consistent with previous research on instrumental sounds and music, which found that participants can differentiate valence and arousal in sounds (Fulcher, 1991; Murray et al., 1996) and recognize discrete emotion categories in music (Juslin, 2013; Juslin &

Laukka, 2003; Scherer, 2004). Interestingly, some physiological studies on facial responses to music revealed a pattern analogous to facial responses to human vocalizations. Studies conducted by Kallinen (2004) and Lundqvist and colleagues (2009) showed increased activity of the *zygomaticus major* in response to the positively valenced music and negatively valenced music triggered the *corrugator supercilii* activity. However, note that the previous studies did not directly compare different kinds of sounds in one experiment. In contrast, our study that directly compared human and instrumental sounds found different facial responses across these stimulus categories.

The fact that spontaneous facial mimicry responses occurred to human vocalizations but not instrumental excerpts may be due to several factors. One possibility is that the process of emotional mimicry serves social functions (Bourgeois & Hess, 2008; Van Baaren et al., 2009). Mimicry is involved in emotion recognition (Oberman et al., 2007; Wood et al., 2016; but see Wołoszyn et al., 2022), it influences the level of sympathy from others, and also regulates relationships (Chartrand & Lakin, 2013; Stel, 2016; Stel et al., 2010). All this may contribute to greater mimicry for a more social stimulus – human vocalizations. Given that Murata and colleagues (2016) showed that the explicit goal to infer the emotional states of the expression producer enhances facial mimicry, we might speculate that such an effect might also occur for cross-channel mimicry of human vocalizations. On the other hand, an instrumental excerpt, as effective as it may be in conveying happiness or fear, does not constitute a social signal of the performer's joy or distress affording a mimicry response. On the other hand, this "social mimicry" explanation does not easily explain why mimicry occurred in an anonymous setting, and, more importantly, why the pattern of results was similar between the sighted and blind participants. After all, blind participants presumably have less experience in producing congruent facial expressions for the audience since they do see and therefore do not mimic others’ facial displays. Therefore, it is worth considering contribution of mimicry as a result of associative learning.

Note that one aspect differentiating vocal and instrumental emotional sounds is the nature of the associative link between auditory emotional stimulus and facial motor activity. In the case of human vocalizations, various processes may underlie auditory-motor associations. One possibility is that the association may be mediated by visual information – in many cases, vocalization is accompanied by the sight of the interaction partner's face, which is spontaneously imitated. Consequently, this may lead to a link between the motor representation of facial expression and the vocalization of a given emotion. Viswanathan and colleagues (2024) propose the opposite direction. They suggest that the development of facial imitation is an associative process that begins with a phase independent of vision. Initially, the infant, as a result of repeated vocalizations, experiences an association between specific acoustic parameters of the sound it produces and the motor activity and proprioceptive signals associated with the sound’s production. The association between the sounds and the view of a face may occur secondarily. Consistent with this proposal that mimicry can be independent of vision are the findings of Hawk, Fischer, and van Kleef (2012). They observed that the mimicry response when listening to a vocalization correlates with the muscle activity present when vocalizing the same emotion. At the neural level, Warren and colleagues (2006) observed an analogous phenomenon – listening to non-verbal vocalizations is associated with the activity of premotor areas involved in facial movements. Therefore, the co-occurrence of auditory and motor elements during one's vocalizations may be responsible for cross-channel mimicry.

More recent studies by Arias and colleagues (2018, 2021), whose results clearly indicate the specificity of human vocal expressions, particularly expressions of happiness, show a slightly different dimension to the relationship between emotional vocalizations and facial response.

They demonstrated that what elicits a mimicry response is a specific acoustic cue that arises from the contraction of the *zygomaticus major* while smiling during sound production. Therefore, facial muscle activity not only impacts how the face looks but also shapes the acoustic features of a sound. Thus, even though musical and vocal expressions have some common acoustic features, which differentiate emotional and neutral sounds, the study by Arias and colleagues (2018) showed that the acoustic cue, which is the consequence of specific oro-facial gesture, is the aspect that plays a decisive role in the mimicry process.

In the context of the differences in facial responses to vocal and instrumental sounds, it is notable that differences in muscles' activity did not simply reflect the self-reported ratings of these stimuli on the valence and arousal dimensions. Specifically, although vocal expressions of happiness were associated with a significantly higher facial activity index than instrumental expressions, participants evaluated them as equally positive. In contrast, sighted participants rated fear vocalizations as more negative and arousing than instrumental fear expressions, which is not consistent with the values of facial activity indexes. This pattern of results indicates that facial responses to auditory displays of emotion are not merely a function of their subjective evaluation, but reflect additional processes, presumably related to associative learning (see also Shaham & Aviezer, 2022) and perhaps inferences about the mental state of the sound producer (Murata et al., 2016).

4.3. The role of vision

In both sighted and blind participants, facial responses depended on the emotion category, and this effect was observed only for human vocalizations. For instrumental sounds, which were rated similarly to vocalizations, neither group showed variation in facial activity due to emotion category. As mentioned earlier, the interaction of group, sound type, and emotion was not significant, suggesting similarities between groups. Follow-up analyses focused only on the blind group confirmed that the blind responded analogously to the sighted group and showed variation in facial response indexes by emotion category, and only for human vocalizations. Overall, the pattern of results suggests the primary importance of auditory-motor association for the development of facial mimicry of auditory stimuli. That is, both the blind and the sighted showed facial responses to emotional sounds, but only if those sounds were vocalizations, whose production is associated with facial movements (Arias et al., 2018).

Still, there was one interesting difference between sighted and blind participants. In blind people, the focused follow-up analyses revealed less differentiation between emotion categories in facial responses to some specific human vocalizations. In this group, the facial activity index (and also the separate activity of the *zygomaticus major*; see Supplementary Materials, Section 1) associated with sad human vocalizations was higher and, contrary to sighted participants, was not significantly different from the index values for the other sound categories, including happy vocalizations. This is consistent with proposals suggesting that visual input and social feedback may play a role in developing more subtle emotional differentiation.

4.4. Temporal dynamics of facial activity

Analysis of the time courses of specific muscle activity, possible with fEMG, shows what lies beneath the averaged responses in the *zygomaticus major* and the *corrugator supercilii*. Although in both groups, the time courses of muscles' activity differed depending on emotion category, in blind individuals the primary difference between the emotions was not the different shape of the reaction as observed in the sighted, but different strength of the increase in and duration of the activity. Again, these differences may reflect the role of visual and social feedback in shaping the dynamics of emotional expressions.

As for the difference in the response dynamics between human vocalizations and instrumental music in the sighted group, the sound type did not statistically significantly modulate the time course of the *zygomatic* muscle to different emotion categories. However, the observed pattern of variation between sound types was similar to that present in blind participants: only in the case of human vocalizations of happiness did the dynamics of the muscle response to happy expressions visibly differ from the courses for the other emotion categories. On the other hand, happy expressions were not associated with a notable "smiling muscle" response for instrumental sounds.

4.5. Visual experience and temporal dynamics of muscle response

When considering the issue of cross-channel mimicry in blind people, it is worth taking into account both aspects of facial activity - average responses and temporal dynamics. On the one hand, the effects in terms of facial activity indexes, a measure combining the mean activity of the *zygomaticus major* and the *corrugator supercilii* muscles, did not differ between the sighted and blind groups, which would indicate an analogous extent of mimicry in both groups. Similarly, the dynamics of the *corrugator supercilii* in response to emotional expressions did not differ across groups.

Interestingly, the groups differed significantly in the response dynamics of the *zygomaticus major.* In sighted people, once a certain level of muscle activity is reached within 2 seconds, we observe a plateau lasting a few seconds and a decrease in the activity of the muscle, whereas in blind people, once a similar level is reached within two seconds, we still observe a relatively constant increase in activity, followed by a decrease. This suggests that visual experiences can modulate the shape of the response of at least some facial muscles. This finding is consistent with the emphasis that mimicry is particularly concerned with positive emotions (Bourgeois & Hess, 2008; Fischer et al., 2012; Hinsz & Tomhave, 1991; Olszanowski et al., 2020) and the argument that the *zygomaticus major* muscle can be considered "more social." As indicated earlier, although the smile is considered by some researchers primarily as an expression of experienced positive emotions (Ekman, 1992, 1993; Ekman & Friesen, 1978), according to the position of behavioral ecology (Fridlund, 1994, 2017), it also has important functions in social interactions. The social-functional approach (Martin et al., 2017) lists three elementary functions of smiling: it can reinforce desirable behavior in the recipient, signal dominance, and serve to sustain relationships by communicating an open attitude and willingness to establish rapport. It is this last function that seems to be most strongly linked to emotional mimicry.

4.6. Possible implications for the debate on the innateness of facial expressions production.

It may be useful to speculate cautiously about the possible implications of our results for the debate about the innateness of facial expression production. Numerous studies on emotion expressions in blind people, both children and adults, reviewed by Valente, Theurell, and Gentaza (2018) suggest that the facial expressions of blind people generally resemble those of sighted people, which suggests at least some innateness in production (Matsumoto et al., 2008; Matsumoto & Willingham, 2009). It is worth noting though that these studies were concerned with expressions either posed or spontaneously occurring in situations of high emotional arousal. Therefore, their conclusions may not necessarily translate to facial responses produced to mild intensity human vocalizations and instrumental sounds. Perhaps more importantly, the key element in the phenomenon in mimicry is the facial movements not only reflect ("express") the perceiver's emotional state, but also reflect copying some element of emotional expressions of others, and modulating one's own expression for others. Furthermore, as we emphasized throughout, the facial response to sounds is at least partially shaped by learning of links between motor production and purely acoustic perception.

In any case, it is possible that visual experience plays a small role in the original shaping of emotional expressions but influences the later and more socially determined process of emotional mimicry, perhaps modulating the facial response on top of any patterns that are produce due to the perceivers' emotional state (Matsumoto & Willingham, 2009). As discussed, our temporal dynamics data on the zygomaticus are generally consistent with this possibility. While our data require caution in interpretation, especially since blind participants varied in terms of the time of vision loss and only some were congenitally blind, our approach is important as it highlights the need to use physiological measures in determining the relevant dynamics and comparing emotion stimuli that vary in terms of associations that underlie automatic mimicry (i.e., vocal and instrumental sound). One exciting, though technically difficult, future study could examine responses to human vocalizations of participants who are both blind and deaf, as they would be lacking the opportunity for both visual-motor as well as acoustic-motor learning of associations between a positive stimulus and a particular facial configuration. If such participants produce robust facial expressions to, say, affective odors, this would suggest a role of innate components.

4.7. Summary

Our study indicates that mimicry responses can operate in a cross-channel manner. Interestingly, we found the most robust facial responses to affect vocalizations, not similarity rated instrumental sounds. The similarity in response to vocalizations between sighted and blind participants suggests that visual experiences might not be necessary for this kind of cross-channel imitation. However, social facial feedback might play a role in shaping more subtle emotional differentiation, including temporal dynamics of the *zygomaticus major*. The observations regarding the dynamics of the muscles' activity also demonstrate the value of the physiological measurements in their ability to capture quick and dynamic changes in spontaneous facial activity.

Declarations

Compliance with Ethical Standards

This study was performed in line with the principles of the Declaration of Helsinki. The design of the study was approved by the Ethics Committee for Experimental Research at the Institute of Psychology, Jagiellonian University in Krakow (decision number KE/04/062018). Informed consent was obtained from all individual participants included in the study.

Competing Interests

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

References

Arias, P., Belin, P., & Aucouturier, J. J. (2018). Auditory smiles trigger unconscious facial imitation. *Current Biology*, *28*(14), R782–R783. https://doi.org/10.1016/j.cub.2018.05.084

Arias, P., Bellmann, C., & Aucouturier, J. J. (2021). Facial mimicry in the congenitally blind. *Current Biology*, *31*(19), R1112–R1114. https://doi.org/10.1016/j.cub.2021.08.059

Arnold, A. J., & Winkielman, P. (2020). The mimicry among us: Intra- and inter-personal mechanisms of spontaneous mimicry. *Journal of Nonverbal Behavior*, *44*(1), 195–212. https://doi.org/10.1007/s10919-019-00324-z

Ashton-James, C., Van Baaren, R. B., Chartrand, T. L., Decety, J., & Karremans, J. (2007). Mimicry and me: The impact of mimicry on self-construal. *Social Cognition*, *25*(4), 518–535. https://doi.org/10.1521/soco.2007.25.4.518

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expressions. *Journal* *of Personality and Social Psychology*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I show how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, *50*(2), 322–329. https://doi.org/10.1037//0022-3514.50.2.322

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. https://doi.org/10.3758/BRM.40.2.531

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 401–432). Cambridge University Press.

Blairy, S., & Hess, U. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*(2), 129–141. https://doi.org/10.1016/s0167-8760(00)00161-6

Bornemann, B., Winkielman, P., & der Meer, E. van. (2012). Can you feel what you do not see? Using internal feedback to detect briefly presented emotional stimuli. *International Journal of Psychophysiology*, *85*(1), 116–124. https://doi.org/10.1016/j.ijpsycho.2011.04.007

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*(3), 343–352. https://doi.org/10.1016/j.biopsycho.2007.11.008

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*(2), 204–215. https://doi.org/10.1017/S0048577200990012

Briefer, E. F. (2018). Vocal contagion of emotions in non-human animals. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1873). https://doi.org/10.1098/rspb.2017.2783

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: The role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2369–2380. https://doi.org/10.1098/rstb.2009.0048

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*, 285–308. https://doi.org/10.1146/annurev-psych-113011-143754

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature*, *589*(7841), 251–257. https://doi.org/10.1038/s41586-020-3037-7

Darwin, C. (2009). *The expression of the emotions in man and animals*. Cambridge University Press.

Davis, J. D., Coulson, S., Blaison, C., Hess, U., & Winkielman, P. (2022). Mimicry of partially occluded emotional faces: Do we mimic what we see or what we know? *Cognition and Emotion*, *36*(8), 1555–1575.

https://doi.org/10.1080/02699931.2022.2135490

Davis, J. D., Winkielman, P., & Coulson, S. (2017). Sensorimotor simulation and emotion processing: Impairing facial action increases semantic retrieval demands. *Cognitive, Affective and Behavioral Neuroscience*, *17*(3), 652–664. https://doi.org/10.3758/s13415-017-0503-2

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology*, *19*(6), 643–647. https://doi.org/10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U. (1990). Facial electromyographic reactions and autonomic activity to auditory stimuli. *Biological Psychology*, *31*(2), 137–147. https://doi.org/10.1016/0301-0511(90)90013-M

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. https://doi.org/10.1111/1467-9280.00221

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–5), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384–392.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: Research nexus*. Network Research Information.

Fischer, A. H., Becker, D., & Veenstra, L. (2012). Emotional mimicry in social context: The case of disgust and pride. *Frontiers in Psychology*, *3*(475), 1–9. https://doi.org/10.3389/fpsyg.2012.00475

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. Academic Press.

Fridlund, A. J. (2017). The behavioral ecology view of facial displays, 25 years later. In J. M.

Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expressions* (pp. 77–92). Oxford University Press.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*(5), 567–589.

Fulcher, J. A. (1991). Vocal affect expression as an indicator of affective response. *Behavior Research Methods, Instruments, & Computers*, *23*(2), 306–313. https://doi.org/10.3758/BF03203384

Gallucci, M. (2019). *GAMLj: General analyses for linear models* [Computer software]. jamovi module. https://gamlj.github.io/

Hawk, S. T., & Fischer, A. H. (2016). More than just a mirror: Examining the cross-channel mimicry of emotional expressions. In U. Hess & A. H. Fischer (Eds.), *Emotional Mimicry in Social Context* (pp. 107–124). Cambridge University Press.

Hawk, S. T., Fischer, A. H., & Van Kleef, G. A. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social Psychology*, *102*(4), 796–814. https://doi.org/10.1037/a0026234

Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). “Worth a thousand words”: Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, *9*(3), 293–305. https://doi.org/10.1037/a0015178

Hess, U., Arslan, R., Mauersberger, H., Blaison, C., Dufner, M., Denissen, J. J. A., & Ziegler, M. (2017). Reliability of surface facial electromyography. *Psychophysiology*, *54*(1), 12–23. https://doi.org/10.1111/psyp.12676

Hess, U., & Fischer, A. H. (2013). Emotional mimicry as social regulation. *Personality and Social Psychology Review*, *17*(2), 142–157. https://doi.org/10.1177/1088868312472607

Hess, U., & Fischer, A. H. (2014). Emotional mimicry: Why and when we mimic emotions.

Social and Personality Psychology Compass, *8*(2), 45–57. https://doi.org/10.1111/spc3.12083

Hinsz, V. B., & Tomhave, J. A. (1991). Smile and (half) world smiles with you, frown and you frown alone. *Personality and Social Psychology Bulletin*, *17*(5), 586–592. https://doi.org/10.1177/0146167291175014

Iturregui-Gallardo, G., & Méndez-Ulrich, J. L. (2020). Towards the creation of a tactile version of the Self-Assessment Manikin (T-SAM) for the emotional assessment of visually impaired people. *International Journal of Disability, Development and Education*, *67*(6), 657–674. https://doi.org/10.1080/1034912X.2019.1626007

Izard, C. E. (1979). *The Maximally Discriminative Facial Movement Coding System*. Instructional Resource Center, University of Delaware.

Juslin, P. N. (2013). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, *4*(596). https://doi.org/10.3389/fpsyg.2013.00596

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kallinen, K. (2004). Emotion-related psychophysiological responses to listening to music with eyes-open versus eyes-closed: Electrodermal (EDA), electrocardiac (ECG), and electromyographic (EMG) measures. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th International Conference on Music Perception and Cognition, Evanston, USA, August 2004* (pp. 299–301). Causal Productions.

Kamiloğlu, R. G., Boateng, G., Balabanova, A., Cao, C., & Sauter, D. A. (2021). Superior communication of positive emotions through nonverbal vocalisations compared to speech prosody. *Journal of Nonverbal Behavior*, *45*(4), 419–454.

https://doi.org/10.1007/s10919-021-00375-1

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, *68*(3), 441–454. https://doi.org/10.1093/acprof:oso/9780195179644.003.0007

Kret, M. E., Stekelenburg, J. J., Roelofs, K., & Gelder, B. D. (2013). *Perception of face and body expressions using electromyography, pupillometry and gaze measures*. *4*(28), 1–12. https://doi.org/10.3389/fpsyg.2013.00028

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Leinonen, L., Hiltunen, T., Linnankoski, I., & Laakso, M.-L. (1997). Expression of emotional–motivational connotations with a one-word utterance. *The Journal of the Acoustical Society of America*, *102*(3), 1853–1863. https://doi.org/10.1121/1.420109

Liu, X., Xu, Y., Alter, K., & Tuomainen, J. (2018). Emotional connotations of musical instrument timbre in comparison with emotional speech prosody: Evidence from acoustics and event-related potentials. *Frontiers in Psychology*, *9*(737), 1–10. https://doi.org/10.3389/fpsyg.2018.00737

Lundqvist, L. O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music*, *37*(1), 61–90. https://doi.org/10.1177/0305735607086048

Lundqvist, L. O., & Dimberg, U. (1995). Facial expressions are contagious. *Journal of Psychophysiology*, *9*(3), 203–211.

Martin, J., Rychlowska, M., Wood, A., & Niedenthal, P. (2017). Smiles as multipurpose social signals. *Trends in Cognitive Sciences*, *21*(11), 864–877.

https://doi.org/10.1016/j.tics.2017.08.007

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. In M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of Emotions* (pp. 211–233). The Guilford Press. https://doi.org/10.1177/0145445598022200l

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, *96*(1), 1–10. https://doi.org/10.1037/a0014037

Matsumoto, D., Willingham, B., & Olide, A. (2009). Sequential Dynamics of Culturally Moderated Facial Expressions of Emotion. *Psychological Science*, *20*(10), 1269–1274. https://doi.org/10.1111/j.1467-9280.2009.02438.x

Matsumoto, D., & Wilson, M. (2023). Incorporating Consciousness into an Understanding of Emotion and Nonverbal Behavior. *Emotion Review*, *15*(4), 332–347. https://doi.org/10.1177/17540739231163177

Moody, E. J., McIntosh, D. N., Mann, L. J., & Weisser, K. R. (2007). More than mere mimicry? The influence of emotion on rapid facial reactions to faces. *Emotion*, *7*(2), 447–457. https://doi.org/10.1037/1528-3542.7.2.447

Murata, A., Saito, H., Schug, J., Ogawa, K., & Kameda, T. (2016). Spontaneous facial mimicry is enhanced by the goal of inferring emotional states: Evidence for moderation of "automatic" mimicry by higher cognitive processes. *PLOS One*, *11*(4), e0153128. https://doi.org/10.1371/journal.pone.0153128

Murray, I. R., Baber, C., & South, A. (1996). Towards a definition and working model of stress and its effects on speech. *Speech Communication*, *20*(1–2), 3–12. https://doi.org/10.1016/S0167-6393(96)00040-4

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, *2*(3–4), 167–178. https://doi.org/10.1080/17470910701391943

Olszanowski, M., Wróbel, M., & Hess, U. (2020). Mimicking and sharing emotions: A re-examination of the link between facial mimicry and emotional contagion. *Cognition and Emotion*, *34*(2), 367–376. https://doi.org/10.1080/02699931.2019.1611543

Palagi, E., Celeghin, A., Tamietto, M., Winkielman, P., & Norscia, I. (2020). The neuroethology of spontaneous mimicry and emotional contagion in human and non-human animals. *Neuroscience & Biobehavioral Reviews*, *111*, 149–165. https://doi.org/10.1016/j.neubiorev.2020.01.020

Paquette, S., Peretz, I., & Belin, P. (2013). The “Musical Emotional Bursts”: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*(509), 1–7. https://doi.org/10.3389/fpsyg.2013.00509

Peirce, J. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(10). https://doi.org/10.3389/neuro.11.010.2008

Planalp, S. (1996). Communicating emotion in everyday life: Cues, channels, and processes. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion* (pp. 29–48). Academic Press. https://doi.org/10.1016/b978-0-12-057770-5.x5000-0

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In O. Neumann & W. Prinz (Eds.), *Relationships Between Perception and Action: Current Approaches* (pp. 167–201). Springer. https://doi.org/10.1007/978-3-642-75348-0_7

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

RStudio Team. (2023). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/

Rinn, W. E. (1991). Neuropsychology of facial expression. In R. S. Feldman & B. Rimé

(Eds.), *Fundamentals of nonverbal behavior* (pp. 3–30). Cambridge University Press.

Russell, J. A., & Fernandez-Dols, J. M. (1997). The Psychology of Facial Expression. In *The Psychology of Facial Expression*. Cambridge University Press. https://doi.org/10.1017/cbo9780511659911

Scherer, K. R. (1981). Speech and emotional states. In J. K. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 189–220). Grune and Stratton.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*(2), 143–165. https://doi.org/10.1037/0033-2909.99.2.143

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research*, *33*(3), 239–251. https://doi.org/10.1080/0929821042000317822

Scherer, K. R., Banse, R., Wallbott, H. G., Goldbeck, T., & Klaus, R. (1991). Article vocal cues in emotion encoding and decoding. *Motivation and Emotion*, *15*(2), 123–148. https://doi.org/10.1016/s0167-6393(02)00078-x

Schroder, M. (2003). Experimental study of affect bursts. *Speech Communication*, *40*(1–2), 99–116. https://doi.org/10.1016/s0167-6393(02)00078-x

Shaham, G., & Aviezer, H. (2022). Automatic facial reactions to emotional body expressions are not driven by emotional experience. *Emotion*, *22*(4), 641–652. https://doi.org/10.1037/emo0000742

Snowdon, C. T. (2003). Expression of emotion in nonhuman animals. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 457–480). Oxford University Press.

Stel, M. (2016). The role of mimicry in understanding the emotions of others. In U. Hess, A. Fischer (Eds.). *Emotional Mimicry in Social Context* (pp. 27–43). Cambridge University Press.

Stel, M., Blascovich, J., MacCall, C., Mastop, J., Van Baaren, R. B., & Vonk, R. (2010). Mimicking disliked others: Effects of a priori liking on the mimicry-liking link. *European Journal of Social Psychology*, *40*, 867–880. https://doi.org/10.1002/ejsp.655

Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., & Gelder, B. D. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences*, *106*(42), 17661–17666. https://doi.org/10.1073/pnas.0908994106

The Jamovi project. (2023). *Jamovi* (2.4) [Computer software]. https://www.jamovi.org

Tiedens, L. Z., Ellsworth, P. C., & Mesquita, B. (2000). Sentimental stereotypes: Emotional expectations for high- and low-status group members. *Society for Personality and Social Psychology*, *26*(5), 560–575.

Valente, D., Theurel, A., & Gentaz, E. (2018). The role of visual experience in the production of emotional facial expressions by blind people: A review. *Psychonomic Bulletin and Review*, *25*(2), 483–497. https://doi.org/10.3758/s13423-017-1338-0

Van Baaren, R. B., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2381–2389. https://doi.org/10.1098/rstb.2009.0057

Vilaverde, R. F., Horchak, O. V., Pinheiro, A. P., Scott, S. K., Korb, S., & Lima, C. F. (2024). Inhibiting orofacial mimicry affects authenticity perception in vocal emotions. *Emotion*, *24*(6), 1376–1385. https://doi.org/10.1037/emo0001361

Viswanathan, N. K., de Klerk, C. C. J. M., Wass, S. V., & Goupil, L. (2024). Learning to imitate facial expressions through sound. *Developmental Review*, *73*, 101137. https://doi.org/10.1016/j.dr.2024.101137

Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., Rosen, S., & Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor “mirror” system. *Journal of Neuroscience*, *26*(50), 13067–13075. https://doi.org/10.1523/JNEUROSCI.3907-06.2006

Weisbuch, M., & Adams Jr., R. B. (2012). The functional forecast model of emotion expression processing. *Social and Personality Psychology Compass*, *6*(7), 499–514. https://doi.org/10.1111/j.1751-9004.2012.00443.x

Wołoszyn, K., Hohol, M., Kuniecki, M., & Winkielman, P. (2022). Restricting movements of lower face leaves recognition of emotional vocalizations intact but introduces a valence positivity bias. *Scientific Reports*, *12*(1), 16101. https://doi.org/10.1038/s41598-022-18888-0

Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. M. (2016). Fashioning the face: Sensorimotor simulation contributes to facial expression recognition. *Trends in Cognitive Sciences*, *20*(3), 227–240. https://doi.org/10.1016/j.tics.2015.12.010

Supplementary Materials

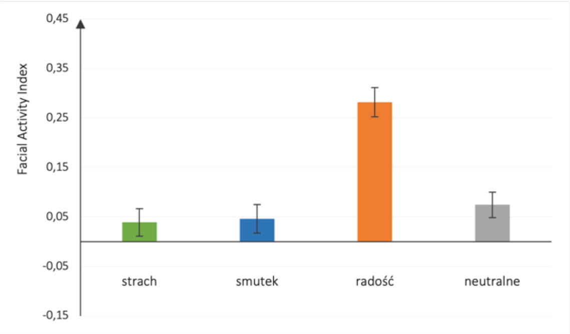
to

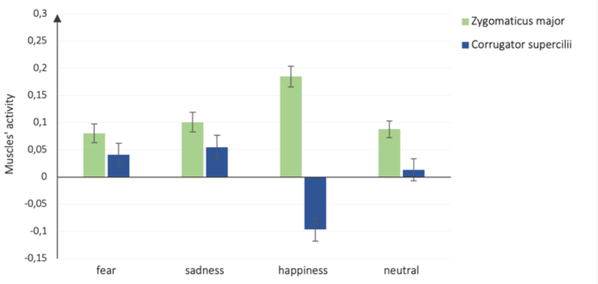
Facing emotional vocalizations and instrumental sounds: Sighted and blind individuals spontaneously and selectively activate facial muscles in response to emotional stimuli

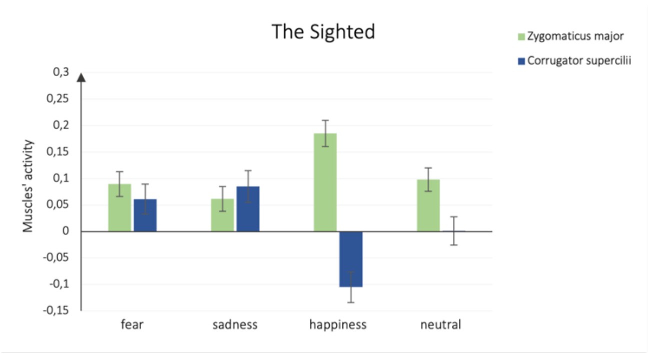
Kinga Wołoszyn, Mateusz Hohol, Michał Kuniecki, Piotr Winkielman

1. Facial activity indexes and individual muscles' activity

Facial activity indexes, which are presented in the main body of the manuscript, are obtained by subtracting the standardized activity of the corrugator supercilii from the standardized activity of the zygomaticus major; they do not show the activity of the individual muscles do not allow to observe the activity of the individual muscles. Below, we present the figures showing the facial activity index.







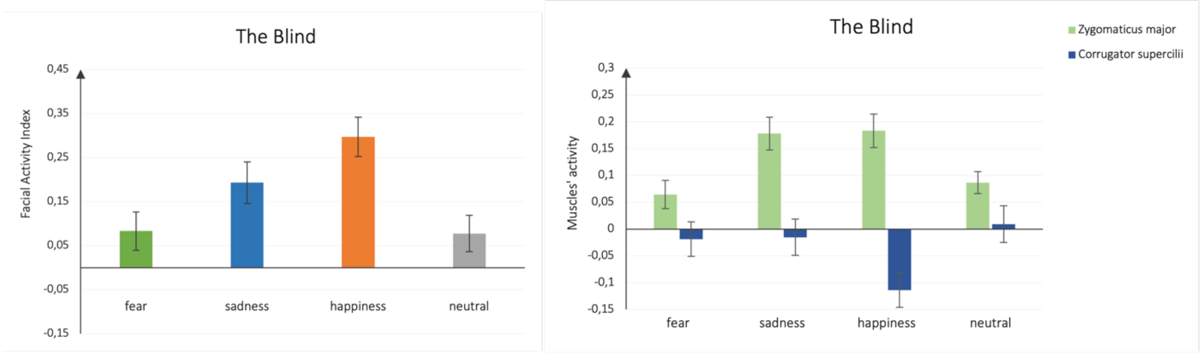


Fig. 1. The facial activity indexes (left side) and the responses of the zygomaticus major and the corrugator supercilii to four emotion categories of sounds.

2. The results of the analyses of the time courses of facial expressions

Table 5. The results of the mixed ANOVA of the time courses of the *zygomaticus major* and the *corrugator supercilii*.

Muscle	Effects	*F*	*df*	*p*	η_p^2
Zygomaticus major	Time	14.46	11, 528	<.001	.232
	Time x Emotion	2.47	33, 1584	<.001	0.05
	Time x Type	4.58	11, 528	<.001	0.09
	Time x Emotion x Type	2.17	33, 1584	.003	0.04
	Group x Time	7.69	11, 528	<.001	0.14
	Group x Time x Emotion	2.01	33, 1584	.006	0.04
	Group x Time x Type	1.74	11, 528	.122	0.035
	Group x Time x Emotion x Type	1.30	33, 1584	.174	0.026
Corrugator supercilii	Time	4.39	11, 528	<.001	0.084
	Time x Emotion	2.93	33, 1584	<.001	0.058
	Time x Type	2.03	11, 528	.038	0.041
	Time x Emotion x Type	1.08	33, 1584	.364	0.022
	Group x Time	1.49	11, 528	.196	0.030
	Group x Time x Emotion	1.04	33, 1584	.414	0.021
	Group x Time x Type	0.343	11, 528	.955	0.007
	Group x Time x Emotion x Type	1.06	33, 1584	.387	0.022

Table 6. The results of the mixed ANOVA of the time courses of the *zygomaticus major* and the *corrugator supercilii* for sighted and blind individuals separately.

Muscle	Effects	Sighted				Blind			
		F	*df*	*p*	partial η^2	*F*	*df*	*p*	η_p^2
Zygomaticus major	Time x Emotion	2.76	33, 1023	<.001	0.08	1.78	33, 561	<.005	0.10
	Time x Type	1.40	11, 341	.234	0.04	4.08	11, 187	.002	0.19
	Time x Emotion x Type	1.40	33, 1023	.079	0.04	1.74	33, 561	.007	0.09
Corrugator supercilii	Time x Emotion	2.89	33, 1023	< .001	0.085	1.49	33, 561	.040	0.08
	Time x Type	1.21	11, 341	.292	0.038	1.16	11, 187	.334	0.06
	Time x Emotion x Type	1.37	33, 1023	.081	0.042	0.84	33, 561	.717	0.05

3. The plots presenting the time courses of facial expressions including the CI

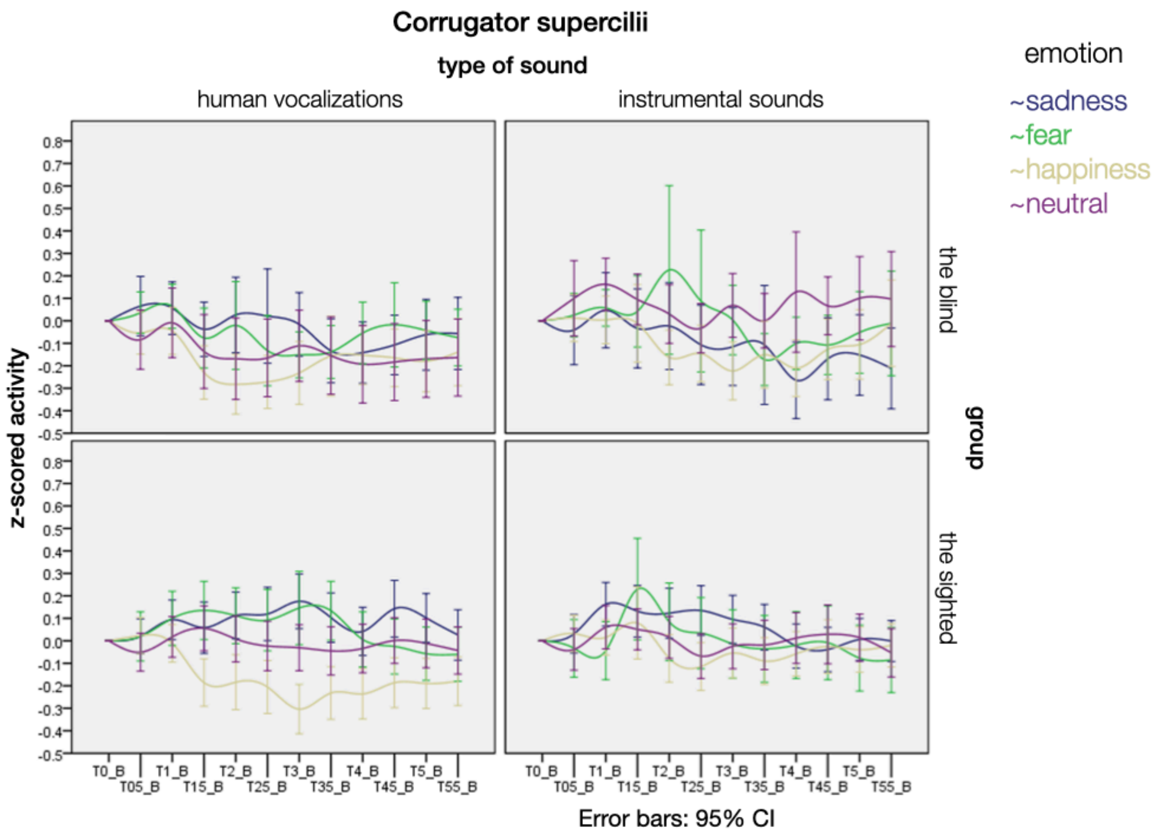


Fig. 2. The time courses of the activity of the corrugator supercilii for two types of sounds and two groups separately.

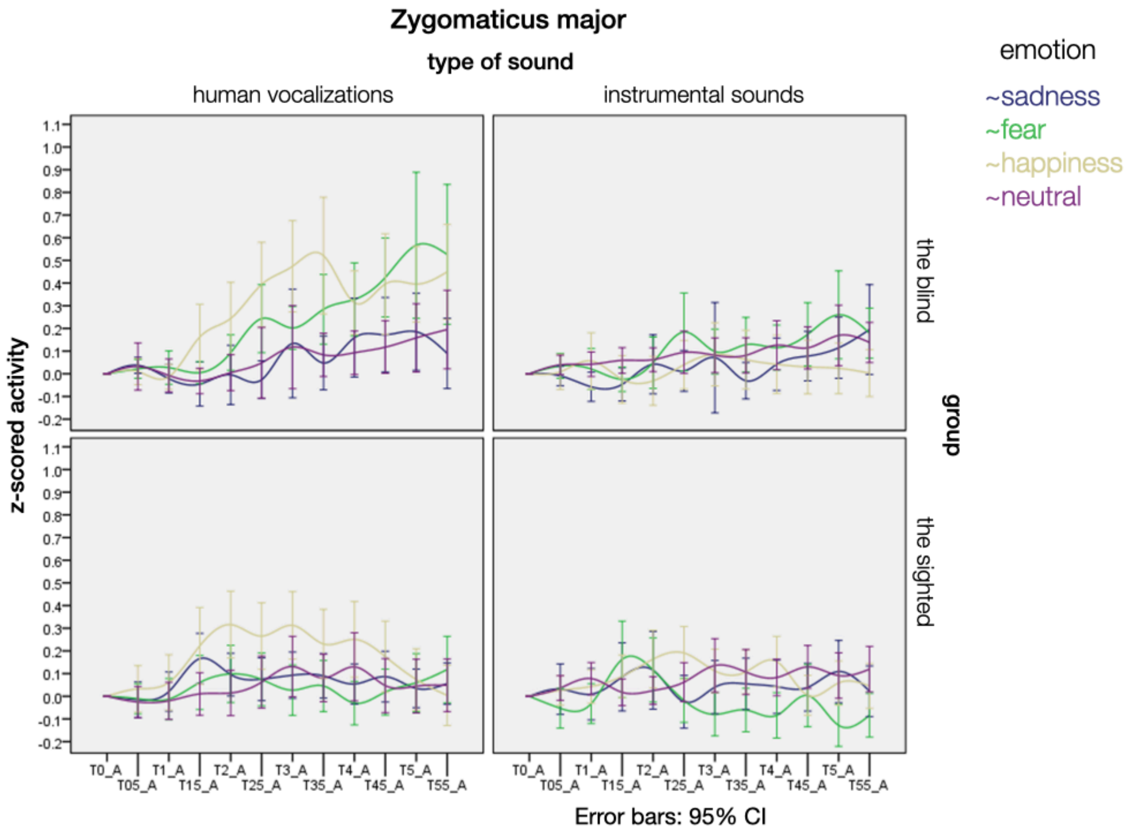


Fig. 3. The time courses of the activity of the zygomaticus major for two types of sounds and two groups separately.

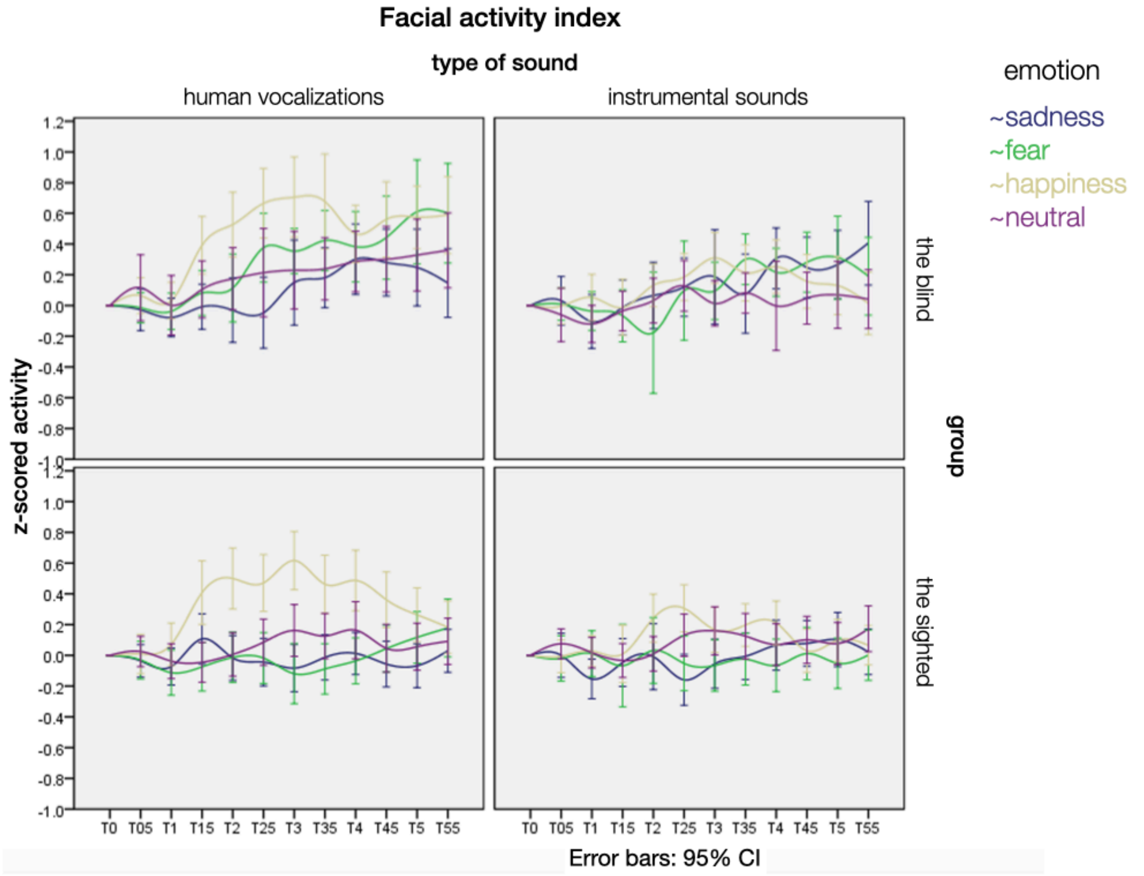


Fig. 4. The time courses of facial responses presented as a facial activity index.

4. Stimuli evaluation

In the second procedure, participants evaluated the stimuli in terms of valence and arousal on a 9-point Likert-type scale (Self-Assessment Manikin; Bradley and Lang, 1994). The scale rages from the most unpleasant/not arousing at all (1) to the most pleasant/most arousing (9). To analyze the impact of emotion and the type of sound on the ratings, we conducted repeated measures ANOVAs with within-subject factors of sound types (vocal, instrumental) and emotion (fear, sadness, happiness, neutral) for two groups separately. To compare the groups, we conducted mixed ANOVAs.

Regarding the valence, in both groups, there was a significant main effect of emotion (sighted: $F(3,93) = 160.153$, $p < .001$; blind: $F(3,51) = 28.674$; $p < .001$). Sighted participants rated sounds expressing happiness as more pleasant ($M = 6.79$, $SE = .127$) than those expressing fear ($M = 3.8$, $SE = .160$), sadness ($M = 3.35$, $SE = .140$), and the neutral sounds ($M = 4.82$, $SE =$

.080). Sad and fearful sounds were rated as more unpleasant than neutral sounds (both differences *p* < .001); happy sounds were rated as more pleasant than neutral ones, and sounds expressing sadness were rated lower than fearful stimuli (*p* = .016). In the blind group, analogously, happy sounds were rated as more pleasant (*M* = 6.32, *SE* = .324) than fearful (*M* = 3.67 *SE* = .275), sad (*M* = 4.03, *SE* = .211) and neutral (*M* = 4.54, *SE* = .244) stimuli (all differences *p* < .001). Moreover, the ratings of fearful sounds were lower than those of the neutral ones (*p* = .048).

In general, the average ratings in term of valence did not vary between the groups significantly (*p* = .771). There was, however, an interaction of group and emotion (*F*(3,144) = 4.90, *p* = .003). Blind participants rated sad sounds as less unpleasant than sighted participants did (4.03 vs. 3.35; see Table 1). Furthermore, in both groups, the type of sound influenced the ratings, with instrumental stimuli rated as more pleasant than human vocalizations (sighted: *F*(1,31) = 20.919, *p* < .001; blind: *F*(1,17) = 6.822; *p* < .018; see Table 2). The groups did not differ in this matter (*p* = .379). The analysis also yielded a significant interaction of type and emotion in both groups (sighted: *F*(3,93) = 24.654, *p* < .001; blind: *F*(3,51) = 8.324, *p* < .001), and a significant difference between the groups (*F*(3,144) = 15.945, *p* < .001). Blind participants rated instrumental sad (*M* = 4.77, *SE* =.284) and happy sounds (*M* = 6.72, *SE* = .273) as more pleasant than sad (*M* = 3.29, *SE* = .291) and happy (*M* = 5.92, *SE* = .424) human vocalizations, while sighted participants rated higher instrumental sounds expressing fear (*M* = 4.92, *SE* = .259 vs. *M* = 2.68, *SE* = .147) and neutral sounds (*M* = 5.1, *SE* = .137 vs. *M* = 4.28, *SE* = .096).

In terms of arousal, we conducted the same analyses as for the valence ratings. First, blind participants rated stimuli as generally less arousing (*M* = 4.85, *SE* = .327) than did the sighted participants (*M* = 5.59, *SE* = .129; *F*(1,48) = 6.213, *p* <.001; see Table 3). In both groups, the emotion category impacted the ratings (sighted: *F*(3,93) = 29.719; *p* < .001; blind: *F*(3,51) = 12.785; *p* < .001). For sighted participants, neutral sounds (*M* = 4.87, *SE* = .131) appeared the least arousing, compared to fearful (*M* = 5.55, *SE* = .177, *p* = .002), sad (*M* = 6.43, *SE* = .168, *p* < .001), and happy stimuli (*M* = 5.51, *SE* = .177, *p* = .008). Sad sounds were rated as the most arousing (all differences *p* < .001). Blind participants, analogously, rated neutral stimuli (*M* = 3.84, *SE* = 459) as less arousing than fearful (*M* = 5.39, *SE* = .292), sad (*M* = 4.99, *SE* = .363), and happy (*M* = 5.16, *SE* = .345) sounds (all differences *p* < .001). The ratings did not differ between the three emotion categories. There was a significant difference between the way both groups rated arousal triggered by the four emotion categories (*F*(1,144) = 7.861, *p* < .001). For the sighted, sounds expressing sadness and neutral sounds appeared more arousing than for the blind (*p* < .001 and *p* = .010, respectively).

In no group did sound type influence the arousal ratings (sighted: *p* = .085; blind: *p* = .74). In both groups, however, there was an interaction between sound type and emotion (sighted: *F*(3,93) = 2.275; *p* < .001; blind: *F*(3,51) = 2.852; *p* = .046), also the groups differed in this term (*F*(3,144) = 6.167, *p* < .001; see Table 4). In the sighted, fearful vocalizations (*M* = 5.94, *SE* = .170) were rated as more arousing than instrumental sounds (*M* = 5.15, *SE* = .229; *p* < .001). There were no significant differences in other emotion categories, neither were there any differences between sound types in the blind.

Table 1. The comparisons of valence ratings of vocal and instrumental stimuli between the sighted and the blind.

Sound type		sighted		blind		difference	
		Mean	SD	Mean	SD	*t*	*p*
vocal							
	Fear	2.68	0.83	3.83	1.42	3.15	.004
	Sad	3.24	0.8	3.29	1.23	0.159	.87
	Happy	6.67	1.28	5.92	1.8	1.72	.092
	Neutral	4.53	.86	4.39	1.39	0.38	.71
instrumental							
	Fear	4.92	1.46	3.51	1.25	3.45	.001
	Sad	3.46	1.02	4.78	1.2	4.1	<.001
	Happy	6.91	0.89	6.72	1.16	0.66	.51
	Neutral	5.11	0.71	4.69	0.89	1.85	.071

Table 2. The comparisons of valence ratings between vocal and instrumental stimuli in the sighted and in the blind.

Group		vocal		instrumental		difference	
		Mean	SD	Mean	SD	*t*	*p*
sighted							
	Fear	2.68	0.83	4.92	1.46	-2.24	<.001
	Sad	3.24	0.8	3.46	1.02	-.22	.19
	Happy	6.67	1.28	6.91	0.89	-.24	.425
	Neutral	4.53	.86	5.11	0.71	-.59	.015
blind							
	Fear	3.83	1.42	3.51	1.25	.33	.3
	Sad	3.29	1.23	4.78	1.2	-1.48	<.001
	Happy	5.92	1.8	6.72	1.16	-.8	.015
	Neutral	4.39	1.39	4.69	0.89	-.3	.25

Table 3. The comparisons of arousal ratings of vocal and instrumental stimuli between the sighted and the blind.

Sound type		sighted		blind		difference	
		Mean	SD	Mean	SD	*t*	*p*
vocal							
	Fear	5.94	.170	5.31	.289	2.01	.051
	Sad	6.39	.173	5.26	.370	3.16	.003
	Happy	5.56	.206	4.99	.374	1.43	.158
	Neutral	4.96	.158	3.72	.467	2.52	.02
instrumental							
	Fear	5.15	.229	5.47	.315	.826	.413
	Sad	6.48	.188	4.72	.420	3.80	< .001
	Happy	5.47	.222	5.33	.382	.322	.749
	Neutral	4.78	.149	3.97	.467	1.65	.113

Table 4. The comparisons of arousal ratings between vocal and instrumental stimuli in the sighted and in the blind.

Group		vocal		instrumental		difference	
		Mean	SD	Mean	SD	*t*	*p*
sighted							
	Fear	5.94	.170	5.15	.229	.784	<.001
	Sad	6.39	.173	6.48	.188	-.083	.53
	Happy	5.56	.206	5.47	.222	.089	.71
	Neutral	4.96	.158	4.78	.149	.182	.26
blind							
	Fear	5.31	.289	5.47	.315	-.162	.46
	Sad	5.26	.370	4.72	.420	.531	.11
	Happy	4.99	.374	5.33	.382	-.340	.29
	Neutral	3.72	.467	3.97	.467	-.250	.18

5. The analysis of facial indexes including valence and arousal ratings.

Table 5 presents the results of the analysis of facial indexes using the LMM model with Facial activity index as a depended variable and Emotion (fear, sad, happy, neutral), Sound_type (vocal, instrumental), Group (sighted, blind), SAM_arousal and SAM_valence as well as their most important interactions as fixed factors and intercept fit across participants and stimuli as random factors. The formula for this analysis in GAMLj is following:

index ~ 1 + Sound_type + Emotion + Group + SAM_valence + SAM_arousal + Sound_type:Emotion + Sound_type:Group + Emotion:Group + Emotion:SAM_valence + Emotion:SAM_arousal + Sound_type:SAM_valence + Sound_type:SAM_arousal + Group:SAM_valence + Group:SAM_arousal + Sound_type:Emotion:Group + Sound_type:Emotion:SAM_valence + Sound_type:Emotion:SAM_arousal + Sound_type:SAM_arousal:Group + Emotion:SAM_arousal:Group + SAM_valence:Soun_type:Group + SAM_valence:Emotion:Group + Sound_type:Emotion:SAM_arousaul:Group + SAM_valence:Sound_type:Emotion:Group + (1 | participant)+(1 | stimulus)

Table 5. Fixed Effect Omnibus tests for analysis including Valence and Arousal ratings.

	F	*df*		*p*
Sound_type	9.1891	1	185.0	0.003
Emotion	1.2132	3	181.1	0.306
Group	6.3243	1	66.9	0.014
SAM_valence	10.0853	1	2177.9	0.002
SAM_arousal	6.2490	1	2784.7	0.012
Group✻ Emotion	1.3146	3	178.0	0.271
Sound_type✻ Group	0.9941	1	3387.0	0.319
Emotion✻Group	2.4757	3	3391.8	0.060
Emotion✻ SAM_valence	2.5182	3	3408.4	0.056
emotion✻ SAM_arousaul	0.2914	3	3373.7	0.832
Sound_type ✻ SAM_valence	0.0217	1	3387.6	0.883
Sound_type ✻ SAM_arousal	0.7375	1	3408.4	0.391
Group ✻ SAM_valence	1.9433	1	2036.2	0.163
Group ✻ SAM_arousaul	2.5296	1	2728.8	0.112
Sound_type✻ Emotion✻ Group	0.1192	3	3381.9	0.949
Sound_type ✻ Emotion✻ SAM_valence	1.0899	3	3401.6	0.352
Sound_type ✻ Emotion ✻ SAM_arousal	0.2651	3	3390.1	0.851
Sound_type ✻ Group ✻ SAM_arousal	7.3611	1	3385.1	0.007
Emotion ✻ Group ✻ SAM_arousal	0.7996	3	3349.4	0.494
Sound_type ✻ Group ✻ SAM_valence	0.9503	1	3167.2	0.330
Emotion ✻ Group ✻ SAM_valence	1.0673	3	3192.1	0.362
Sound_type ✻ Emotion ✻ Group ✻ SAM_arousal	0.0687	3	3373.1	0.977
Sound_type ✻ Emotion ✻ Group ✻ SAM_valence	1.4797	3	3204.6	0.218

Note. Satterthwaite method for degrees of freedom

Reference:

Bradley, M., & Lang, P. J. (1994). Self-Assessment Manikin (SAM). *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9