

Physiological Responses to Aversive and Non-aversive Audiovisual, Audio, and Visual Stimuliaversive and non-aversive audiovisual, audio, and visual stimuli

Sewon Oh, Xuan Yang, William M. Hayes, Ashley Anderson, Douglas H. Wedell, Svetlana V. Shinkareva



PII: S0301-0511(25)00012-2

DOI: <https://doi.org/10.1016/j.biopsycho.2025.108994>

Reference: BIOPSY108994

To appear in: *Biological Psychology*

Received date: 28 June 2024

Revised date: 15 January 2025

Accepted date: 19 January 2025

Please cite this article as: Sewon Oh, Xuan Yang, William M. Hayes, Ashley Anderson, Douglas H. Wedell and Svetlana V. Shinkareva, Physiological Responses to Aversive and Non-aversive Audiovisual, Audio, and Visual Stimuliaversive and non-aversive audiovisual, audio, and visual stimuli, *Biological Psychology*, (2025)  
doi:<https://doi.org/10.1016/j.biopsycho.2025.108994>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Running Head: aversive and non-aversive audiovisual, audio, and visual stimuli

## Physiological Responses to Aversive and Non-aversive Audiovisual, Audio, and Visual Stimuli

Sewon Oh, Xuan Yang, William M. Hayes, Ashley Anderson, Douglas H. Wedell, Svetlana V. Shinkareva

*Department of Psychology, Institute for Mind and Brain, University of South Carolina, Columbia, SC 29201, USA*

Correspondence should be addressed to Svetlana V. Shinkareva, Department of Psychology, Institute for Mind and Brain, University of South Carolina, Columbia, Barnwell College 224, 1512 Pendleton St., South Carolina, 29208, USA; phone: +1(803) 7776189.

E-mail addresses: sewon@email.sc.edu, xy6@email.sc.edu, wmhayes@iu.edu, aea7@email.sc.edu, wedell@sc.edu, shinkareva@sc.edu

Declarations of competing interests: none

Word count: 8968

Number of Figures: 3

Number of Tables: 1

Appendix: 1

Supplemental material: none

Funding source(s): This work was supported by a grant from the Misophonia Research Fund.

Color should not be used for any figures in print.

### Abstract

We examined differences in physiological responses to aversive and non-aversive naturalistic audiovisual stimuli and their auditory and visual components within the same experiment. We recorded five physiological measures that have been shown to be sensitive to affect: electrocardiogram, electromyography (EMG) for zygomaticus major and corrugator supercilii muscles, electrodermal activity (EDA), and skin temperature. Valence and arousal ratings confirmed that aversive stimuli were more negative in valence and higher in arousal than non-

aversive stimuli. Valence also showed an emotional enhancement effect for cross-modal integration. Both heart rate deceleration and facial EMG potentiation for corrugator supercilii were larger for aversive compared to non-aversive conditions for audiovisual stimuli and their auditory components, even after controlling for arousal. Facial EMG potentiation for zygomaticus major was greater for aversive compared to non-aversive conditions for audiovisual stimuli and EDA was greater for aversive compared to non-aversive conditions for visual stimuli. Neither of these effects remained significant after controlling for arousal. These findings provide a benchmark for examining atypical sensory processing of mundane aversive stimuli for clinical populations.

## **Keywords**

physiology, audiovisual, aversiveness, valence, arousal

## **Introduction**

Sights and sounds we encounter in our daily lives significantly influence our emotions, behavior, and interactions with the world around us. For instance, the harsh noise of a chair scraping across the floor can trigger unpleasant feelings, while the gentle crackle of a fire brings about positive responses. These emotional reactions can be defined by two dimensions: valence, an intrinsic aversiveness or attractiveness of a stimulus which spans from negative to positive, and arousal, an intensity of the emotional response, which ranges from low to high (Russell, 2003). These two core affective dimensions partially reflect appetitive and defensive

motivational circuits, with valence generally indicating either defensive or appetitive motivation and arousal indicating the degree of motivational activation (Bradley et al., 2001; Carver & Harmon-Jones, 2009). Affective signals are inherently multimodal, allowing us to experience emotions in richer, more nuanced ways. The interplay between auditory and visual modalities can influence the affective response. For instance, harsh sounds can amplify feelings of discomfort, regardless of the accompanying visual, such as simultaneously hearing and seeing a chair scraping across the floor. This study aims to explore the physiological correlates of affective signals associated with naturalistic audiovisual clips and their auditory and visual components, which we will refer to as the three modalities: audiovisual, auditory, and visual.

Fluctuations in affective states are accompanied by changes in sympathetic and parasympathetic systems reflected in physiological responses. Heart rate (HR) has been associated with both valence and arousal. It has been shown that initial heart rate deceleration was sensitive to various aversive stimuli, such as pictures (Lang et al., 1993; Bradley et al., 2001; Gomez et al., 2016), sounds (Bradley & Lang, 2000; Coutinho & Cangelosi, 2011), films (Baldaro et al., 2004; Bradley et al., 2001; Palomba et al., 2000) and aversive environments designed to elicit fear and pain (Vowles et al., 2006). Facial electromyography (EMG) measured over the brow (corrugator supercilii) has been shown to increase with negative valence for picture (Cacioppo et al., 1986; Greenwald et al., 1989; Kim & Wedell, 2016; Lang et al., 1993; Larsen et al., 2003; Tan et al., 2012), sound (Bradley & Lang, 2000; Jäncke et al., 1996; Kim et al., 2019; Larsen et al., 2003), word (Larsen et al., 2003) and film (Sato et al., 2020; Sato & Kochiyama, 2022) stimuli. EMG over the cheek (zygomaticus major) has been shown to increase with intensity of both positive and negative valence (Bernat et al., 2006; Bradley & Lang, 2000; Codispoti et al., 2008; Dellacherie et al., 2011; Grewe et al., 2007; Lang et al., 1993; Larsen et

al., 2003). The relationship between facial EMG and arousal ratings has not received much attention. One study identified a positive relationship with zygomaticus major potentiation, but not for corrugator supercilii potentiation for videos (Gjoreski et al., 2022), while another study found no relationship with arousal for picture stimuli (Tan et al., 2012). Heightened electrodermal activity (EDA) has been shown to be sensitive to emotionally arousing picture (Amrhein et al., 2004; Bradley & Lang, 2000; Bradley et al., 2001; Kim & Wedell, 2016; Lang et al., 1993), sound (Bradley & Lang, 2000; Greco et al., 2017), and film (Sato et al., 2020; Sato & Kochiyama, 2022) stimuli. Decreased nose-tip skin temperature (SKT) has been associated with increased arousal for picture (Salazar-López et al., 2015) and film stimuli (Sato et al., 2020; Sato & Kochiyama, 2022). Thus, auditory and visual affective stimuli have both been shown to elicit physiological responses that can be measured using various physiological channels (Bradley & Lang, 2000; Lang & Bradley, 2010).

Most of the aforementioned studies included erotic or comedic (Amrhein et al., 2004; Bradley et al., 2001; Gomez et al., 2016; Salazar-López et al., 2015; Sato & Kochiyama, 2022), gore or accident-related (Amrhein et al., 2004; Bradley et al., 2001; Gjoreski et al., 2022; Gomez et al., 2016; Larsen et al., 2003; Salazar-López et al., 2015; Tan et al., 2012) stimuli that tend to elicit strong affective responses. Although stimuli rated high in affect may result in stronger physiological reactions, these stimuli are not suitable for clinical applications such as examining sensory sensitivity in children with autism. The degree to which affective states elicited by less extreme stimuli that one may encounter in daily life will be reflected in physiological responses is unclear and would be informative to researchers focused on sensory sensitivity.

In the natural environment, auditory and visual affective signals tend to co-occur (De Gelder & Bertelson, 2003). Indirect evidence from valence and arousal ratings and event-related

potentials has suggested that there is an enhancement effect, where affective signals presented from multiple modalities elicit more extreme affective responses than either source presented alone (Gao et al., 2018a, Gao et al., 2018b). However, most physiological studies have focused on one of the two modalities in isolation, with a few notable exceptions. Rosa et al. (2017) used International Affective Picture System (IAPS) images, International Affective Digital Sounds (IADS), and their combinations within the same experiment in a between-participant design to examine EDA differences in valence. Baumgartner et al. (2006) used classical music, IAPS images, and a combination of the two in the same experiment to examine differences in HR, EDA and SKT for happiness, sadness, and fear. They reported increased physiological activity for audiovisual and audio stimuli compared to the picture stimuli. In both experiments that examined multimodal differences, the audiovisual combination was created by combining static images with sounds or music from different sources that were congruent on affect. In a small group design, Edelstein et al. (2013) separated the auditory and visual components of video clips that would elicit emotionally varying responses in participants with and without misophonia, a condition of heightened sensitivity to specific sounds. They compared the skin conductance response and aversiveness ratings between the two groups but did not focus on the modality effect. Thus, physiological responses to audiovisual stimuli and their auditory and visual components from the same source have not yet been examined in a non-clinical population. Our experiment builds on this literature and fills this gap.

In the present study, we examined differences in physiological responses to aversive (negative valence, higher arousal) and non-aversive (positive valence, lower arousal) naturalistic audiovisual stimuli that one may encounter in daily life and their auditory and visual components within the same experiment. These conditions differ on arousal, as aversive stimuli tend to be

more arousing (Garrett & Maddock, 2001; Kuppens et al., 2013; Nitschke et al., 2006).

Understanding physiological responses to innocuous aversive stimuli is important for investigating atypical sensory processing in clinical populations, such as those with Autism Spectrum Disorder (Jung et al., 2021), hyperacusis (Singh & Seidman, 2019), a condition where ordinary sounds are perceived as excessively loud or painful, and misophonia (Swedo et al., 2022), a condition characterized by aberrant emotional response to certain sounds. This study makes two important contributions. First, it examines physiological responses to aversive and non-aversive audiovisual clips and their auditory and visual components within the same study. Second, it uses stimuli that may be encountered in everyday life that hold relevance for clinical applications. We recorded five physiological measures that have been shown to be sensitive to affect: electrocardiogram (ECG), EMG for zygomaticus major and corrugator supercilii muscles, EDA, and SKT and investigated how they change as a function of aversiveness for audiovisual stimuli and their auditory and visual components. We formulated four hypotheses. First, we hypothesized that since facial EMG measures have been shown to be sensitive to valence, they should be sensitive to aversiveness. The EMG signals would be higher for aversive compared to non-aversive stimuli (Bradley & Lang, 2000; Lang et al., 1993; Larsen et al., 2003), although mixed results have been shown for EMG<sub>Z</sub> (Bernat et al., 2006; Codispoti et al., 2008). Second, we expected HR deceleration to be greater and EDA to be higher for the aversive condition. Negative high-arousal stimuli typically elicit larger deceleration of HR (Bradley et al., 2001) and larger EDA responses (Bradley & Lang, 2006; Bradley et al., 2001; Bradley & Lang, 2000) as do losses versus gains in choice experiments (Crone et al., 2004; Hayes & Wedell, 2020). Third, the heightened response to aversive stimuli would occur for all three modalities. Fourth, the physiological and self-report sensitivity to aversiveness should be greatest in the audiovisual

condition given past research demonstrating an enhancement effect in cross-modal integration (Gao et al., 2018a, Gao et al., 2018b).

## Method

### Participants

Physiological data were collected from 75 participants ( $M_{\text{age}} = 20.21$  years,  $SD_{\text{age}} = 2.27$  years,  $\text{Range}_{\text{age}} = 18\text{-}31$  years; 60 Females, 15 Males; 71 right-handed, 3 left-handed, 1 ambidextrous) in exchange for a course credit at the University of South Carolina. Data from 73 participants ( $M_{\text{age}} = 20.26$  years,  $SD_{\text{age}} = 2.28$  years,  $\text{Range}_{\text{age}} = 18\text{-}31$  years; 59 Females, 14 Males; 69 right-handed, 3 left-handed, 1 ambidextrous) were analyzed (see Data Cleaning and Analysis section below). All participants had normal or corrected-to-normal vision and normal hearing. The study was carried out in accordance with the procedures and protocols approved by the University of South Carolina Institutional Review Board (Pro00108272) and all participants signed an informed consent form. Based on an *a priori* power analysis conducted in G\*Power (Version 3.1.9.6; Faul et al., 2007), a sample size of 55 participants is required for an effect size of  $f = .2$  and a power of .90 to detect an interaction of aversiveness and modality at  $\alpha = .05$ .

### Stimuli

Stimuli were selected from an in-house database of 5s naturalistic audiovisual recordings developed for studying auditory sensory sensitivity. Forty stimuli (20 aversive, 20 non-aversive) were selected based on aversiveness ratings in the audiovisual modality. The clip content was similar to what one might encounter in daily life, featuring videos and sounds like a squeaky



chair rolling across the floor for an aversive clip, and the soothing crackle of a fire for a non-aversive clip (Appendix A). Each audiovisual clip was separated into auditory and visual components, resulting in corresponding audio-only and video-only stimuli. The audio track of audio and audiovisual stimuli was normalized to a uniform loudness level of –16 Loudness Unit Full Scale. Further, aversiveness ratings of auditory and visual components were validated across two online experiments<sup>1</sup>.

## Procedure

Upon arrival, each participant was seated at the computer, familiarized with the study and valence-arousal grid to rate stimuli. Participants were asked to attend to the stimuli during the stimulus presentation and rate each stimulus when the grid appeared on the screen in terms of valence (varying from negative to positive) and arousal (varying from low to high). Instructions were both shown on the screen and verbalized by a researcher with the prompt: “You will be presented with either videos with sounds, videos without sounds, or sounds without videos. Rate valence and arousal for each stimulus on the 9 x 9 grid. Valence: how negative or positive. Arousal: how calm or excited.” An experimenter attached the electrodes and calibrated the physiological signal connections. Following the preparation phase, participants were asked to sit still and rest for five minutes, during which the physiological signals could stabilize. To

---

<sup>1</sup> The validation experiments were conducted on a separate set of 49 participants ( $M_{age} = 20.10$  years,  $SD_{age} = 1.28$  years; 40 Females, 9 Males). In these validation experiments, the participants rated each audio-only and video-only stimulus on valence, ranging from negative (-4) to positive (4), arousal, ranging from low (1) to high (9), aversiveness, ranging from very aversive (-9) to very pleasant (+9), and content recognition, either Yes or No. All stimuli were recognized by at least 90% of the participants. The aversiveness ratings were significantly higher for aversive ( $_{aver}$ ) than the non-aversive ( $_{non-aver}$ ) stimuli for auditory component ( $M_{aver - non-aver} = -8.49$ ,  $SD_{aver - non-aver} = 3.48$ ),  $t(48) = -17.06$ ,  $p < .001$ ,  $d = 2.44$ , 95% CI [-9.49, -7.49], and for the visual component ( $M_{aver - non-aver} = -7.61$ ,  $SD_{aver - non-aver} = 3.04$ ),  $t(48) = -17.54$ ,  $p < .001$ ,  $d = 2.51$ , 95% CI [-8.48, -6.73]. There were 120 clips in total: 20 stimuli  $\times$  2 (aversiveness: aversive, non-aversive)  $\times$  3 (modality: audiovisual, auditory component, visual component).

familiarize the participants with the task, the experiment started with a practice session consisting of one stimulus from each of the three modalities that were not part of the test stimuli. E-Prime 3.0 software (Version 3.0.3.60 Psychology Software Tools, Pittsburgh, PA) was used to present the stimuli and record valence and arousal rating responses. Participants were presented with 5s audiovisual, audio, and video clips that differed on aversiveness. The experiment took 40 minutes to complete.

### **Design and task**

The physiology experiment used a 2 (aversiveness: aversive, non-aversive)  $\times$  3 (modality: audiovisual, auditory component, visual component) within-participant design. Each of the three modalities consisted of 40 stimuli (20 aversive, 20 non-aversive) divided into four blocks, each comprising ten trials. The block order was randomized, with the restriction that blocks of the same modality were not presented consecutively to avoid boredom. For each block, ten 20s non-repeating trials (five aversive and five non-aversive) were presented in a pseudo-randomized order with the constraint that no more than three trials of the same aversiveness condition were presented consecutively to avoid habituation effects (Fig. 1A).

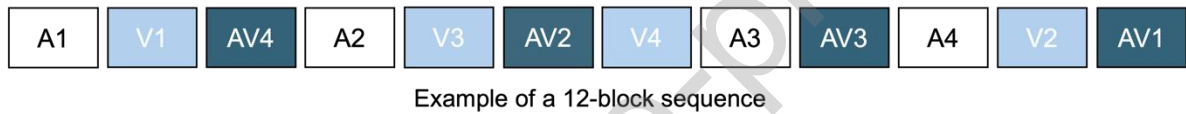
Each 20s trial began with a fixation cross centered on a white screen jittered for 1-2s, followed by a 5s stimulus presentation. During audio trials, an audio icon was presented in the center of the screen while the sound played. Following the stimulus presentation, a fixation cross was presented on a white screen for 10s to allow enough time for physiological signals to return to baseline level. Next, a  $9 \times 9$  two-dimensional affective rating grid appeared for 3-4s, depending on pre-stimulus fixation cross interval to make up 20s for each trial, where participants rated the stimulus on valence (horizontal axis), varying from negative to positive,

and arousal (vertical axis), varying from low to high (Fig. 1B). Simultaneous collection of valence and arousal ratings has been developed by Russell (1989) and has been successfully used in prior studies (e.g., Gao et al., 2018; Kim et al., 2017; Lazar & Pearlman-Avnion, 2014; Scott Killgore, 1998). Failure to respond within the affective rating window resulted in missing data.

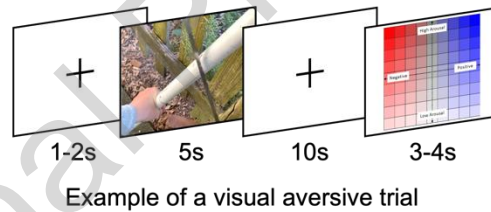
**Figure 1**

*Experiment Procedure and Design*

**A. Experiment consisted of 12-blocks**



**B. Each block consisted of 5 aversive and 5 non-aversive trials from the same modality**



*Note.* **A.** Example of an experimental sequence. The experiment consisted of 12 blocks, four for each modality: audiovisual (AV), auditory (A), and visual (V). Twelve-block sequences were randomized with one constraint: Blocks from the same modality were not presented back-to-back. **B.** Each block consisted of 10 trials from the same modality, five aversive and five non-aversive. Trials were presented in pseudo-randomized order such that no more than three trials of the same aversiveness condition were presented consecutively.

**Physiological Data Acquisition**

Physiological signals were acquired with a BIOPAC MP150 system (BIOPAC Systems Inc., Goleta, CA, USA) and processed with AcqKnowledge 3 software (BIOPAC Systems Inc., Goleta, CA, USA). Five channels were recorded: ECG, EMG for corrugator supercilii and zygomaticus major muscles, EDA, and SKT. Heart rate was measured using an ECG100C amplifier with hypo-allergic electrode gel (GEL100) applied to two disposable electrodes (EL503) placed on the frontal torso, one below the right collar bone and another one below the lowest left rib in the common Lead II configuration. Since concurrently measured EDA provides an automatic grounding, a separate ground connection for ECG was not added. EMG was measured using an EMG100C amplifier with hypo-allergic electrode gel (GEL100) applied to two 4mm reusable Ag/AgCl electrodes (EL654) for corrugator supercilii muscle potentiation (EMG<sub>C</sub>), other two for zygomaticus major muscle potentiation (EMG<sub>Z</sub>), and another one for grounding. The ground electrode was placed in the middle of the inner brows. Adhesive disks (ADD204) were used to attach the reusable snaps onto skin. The guidelines from Fridlund and Cacioppo (1989) were followed to find the appropriate positioning of the EMG<sub>C</sub> and EMG<sub>Z</sub> electrodes. EDA signal was recorded using a GSR100C amplifier with isotonic electrode paste (GEL101A) applied to two disposable Ag/AgCl electrodes (EL507A; 16mm gel cavity diameter, 1.5mm gel cavity depth, 11mm contact area diameter) attached to the left hand's medial phalanges of the index and middle fingers. SKT was measured using a SKT100C amplifier on the palm side of the thumb of the left hand to record skin temperature in Celsius (°C). The sampling rate for all the physiological measures was 1000Hz.

### **Data Preprocessing**

The analysis time window was 0-7s as determined by optimizing physiological effects in another experiment conducted in our lab, that had the same presentation sequence. We kept the analysis window consistent across channels. For ECG, R peaks were automatically detected using a QRS complex detection algorithm in AcqKnowledge 3 (BIOPAC Systems, Inc.). Artifacts were visually identified and manually corrected. Connect Endpoints function in AcqKnowledge 3 (BIOPAC Systems, Inc.) was used for correcting artifact segments and label was inserted or removed for any missing QRS peaks or overestimated peaks, respectively. Heart rate was estimated with beats per minute (BPM) during the 7s post-stimulus onset, and baseline correction was conducted by subtracting mean BPM during the 1s window pre-stimulus onset from mean BPM during the 7s window post-stimulus onset.

The EMG data was first filtered with Band Stop filter configured with a line frequency setting at 60Hz and a Q factor of 5.0 to remove the power line interference, using an infinite impulse response. Next, Band Pass filter was applied with low-frequency cut off at 20Hz and high-frequency cut off at 500Hz to attenuate low-frequency motion artifacts and high-frequency noises, using finite impulse response. The filtered signals were further smoothed with a 250ms moving window to denoise inherent variability of muscle activity and then rectified by taking the absolute value transformation. All the processes were applied to the entire waveform. EMG in response to stimulus was measured as the area under the curve (AUC) calculated for the 7s time window after and 1s time window before the stimulus onset. Log transformation,  $\log_{10}(x)$ , was then applied to both time windows to reduce the skewness of the distributions. Baseline correction was conducted using the log-transformed AUC to account for participants with higher- or lower-than-average signal.

For EDA, phasic EDA was derived online from the tonic EDA with a 0.05 Hz high-pass filter. Both the phasic and tonic EDA were smoothed with a 250ms moving window to reduce noise. Full wave rectification was applied to the smoothed phasic EDA. The AUC calculation, log transformation, and baseline correction were done the same way as in the EMG analysis. For SKT, the skin temperature was averaged across the 7s post-stimulus onset and baseline corrected for the mean of 1s pre-stimulus onset. The signals for all channels were then z-transformed within-participant to allow aggregation of data accounting for individual differences.

### **Data Cleaning and Analysis**

Self-report data analysis and data from EDA and SKT channels included data from 73 participants. The final data set for EMG<sub>Z</sub> analysis included data from 71 participants, EMG<sub>C</sub> and ECG analyses included data from 72 participants.

EMG<sub>Z</sub> analysis excluded two participants because of inadequate signal acquisition due to facial hair interference and cheek biting. EMG<sub>C</sub> analysis excluded one participant because of signal loss. ECG analysis also excluded one participant because of signal loss.

Signals with movement-related artifacts (e.g., yawning; 3.92%) were excluded from the physiological data analysis. Missing responses (0.27%) and responses under 200ms (3.66%) were excluded from the self-report data analysis (Whelan, 2008). Data from two additional participants were excluded from both physiological and self-report data analyses due to a large number of invalid responses (29.17% and 50.00%). Responses to audiovisual stimuli that were not manipulated based on individuals' ratings (3.21%) were excluded from both the physiological and self-report data analyses. Further, signals outside 3SD from the mean within individuals were excluded from the physiological data analysis. As a result, 5.79% of EDA,

8.01% of EMG<sub>C</sub>, 10.14% of EMG<sub>Z</sub>, 6.04% of SKT, 7.16% of ECG, and 6.61% of valence and arousal trials across participants were excluded.

A 2 (aversiveness: aversive, non-aversive) x 3 (modality: A, V, AV) repeated-measures ANOVA was used to analyze the data for each channel. Greenhouse-Geiser adjustment was used when the sphericity assumption did not hold (Mauchly's test). The results were evaluated at the .05 level of statistical significance. We report the uncorrected *df* and the corrected *p*-values where applicable. This was followed by planned comparisons to test the simple main effects assessing signal difference to aversiveness type within each modality and rating. *P*-values for planned *t*-tests were not adjusted for multiple comparisons. The same analyses were conducted using residualized physiological signals to assess the effects of aversiveness controlling for arousal. Physiological signals adjusted for arousal were first extracted by modeling the dependent variables from arousal ratings only. Then the residuals were used for the ANOVA. The results were evaluated at the .05 level of statistical significance.

We conducted an epoch analysis to closely examine the temporal structure of aversiveness differences across physiological channels. We applied the same preprocessing pipeline to extract z-transformed data for each 1-second epoch from the onset of stimulus presentation to 7 seconds after the stimulus onset, resulting in time-series data for each trial and participant. We used the MNE-Python package to identify continuous temporal epochs that significantly differ between aversive and non-aversive conditions for each physiological channel. The statistical significance of clusters was assessed using a permutation test. For each channel, we repeated the epoch analysis 1,000 times with shuffled physiological time series, generating a null distribution of 1,000 values under the null hypothesis that no significant difference exists between aversive and non-aversive conditions.

## Results

### *Valence and Arousal Ratings*

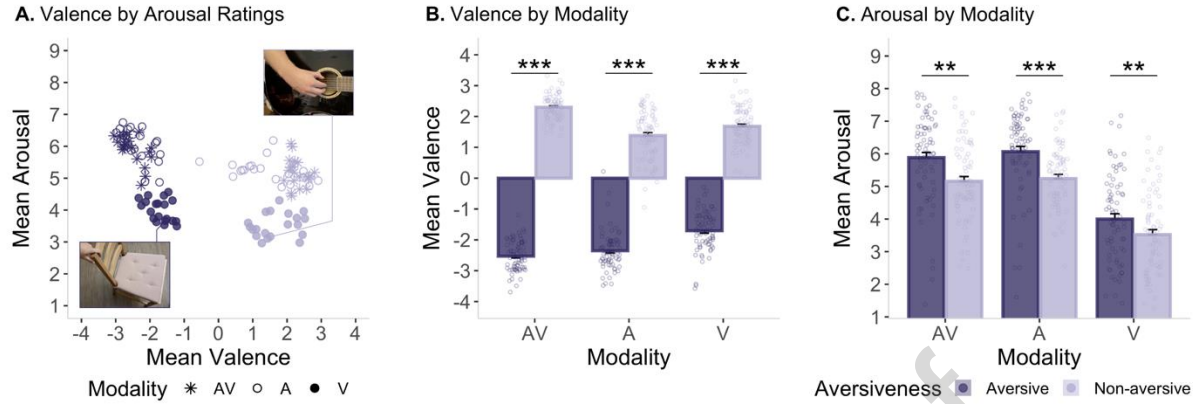
A 2 x 3 repeated measures ANOVA was performed on affect ratings to confirm the experimental manipulation. There were main effects of aversiveness, modality, and interaction on valence ratings. The valence ratings confirmed the experimental manipulation, with more negative ratings for aversive compared to non-aversive conditions for audiovisual ( $M_{aver} = -2.53$ ,  $SD_{aver} = 0.48$ ;  $M_{non-aver} = 2.30$ ,  $SD_{non-aver} = 0.40$ ), auditory ( $M_{aver} = -2.35$ ,  $SD_{aver} = 0.61$ ,  $M_{non-aver} = 1.38$ ,  $SD_{non-aver} = 0.80$ ), and visual ( $M_{aver} = -1.71$ ,  $SD_{aver} = 0.66$ ;  $M_{non-aver} = 1.68$ ,  $SD_{non-aver} = 0.62$ ) modalities (Table 1, Fig. 2A, 2B). Interaction was consistent with enhancement effect in which valence differences for AV stimuli were greater than those for A and V components.

There were main effects of aversiveness, modality, and interaction on arousal ratings. Arousal was significantly higher for aversive compared to non-aversive conditions for audiovisual ( $M_{aver} = 5.88$ ,  $SD_{aver} = 1.34$ ;  $M_{non-aver} = 5.16$ ,  $SD_{non-aver} = 1.22$ ), auditory ( $M_{aver} = 6.07$ ,  $SD_{aver} = 1.37$ ;  $M_{non-aver} = 5.24$ ,  $SD_{non-aver} = 1.09$ ), and visual ( $M_{aver} = 4.00$ ,  $SD_{aver} = 1.39$ ;  $M_{non-aver} = 3.52$ ,  $SD_{non-aver} = 1.30$ ) modalities (Table 1, Fig. 2A, 2C). Interaction was driven by greater aversiveness differences for AV stimuli and A component compared to V component.

### **Figure 2**

#### *Valence and Arousal Ratings across all Conditions*





Note. **A.** Auditory, visual, and audiovisual stimuli in valence-by-arousal space. **B.** Valence ratings. Aversive stimuli were rated significantly more negative than non-aversive stimuli for all modalities. **C.** Arousal ratings. Aversive stimuli were rated significantly more arousing than non-aversive stimuli for all modalities. The visual component was rated as least arousing. Only the simple main effects of aversiveness type are visualized. \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ .

### Heart Rate

We examined differences in HR for aversive and non-aversive conditions across the three modalities. There was an interaction effect of aversiveness and modality (Table 1). Planned comparisons were conducted to examine the aversiveness effect for each modality. Consistent with an enhancement effect, HR was significantly lower for aversive compared to non-aversive conditions for audiovisual ( $M_{\text{aver}} = -0.04$ ,  $SD_{\text{aver}} = 0.23$ ,  $M_{\text{non-aver}} = 0.15$ ,  $SD_{\text{non-aver}} = 0.23$ ) and auditory ( $M_{\text{aver}} = 0.03$ ,  $SD_{\text{aver}} = 0.21$ ,  $M_{\text{non-aver}} = 0.12$ ,  $SD_{\text{non-aver}} = 0.21$ ), but not for visual ( $M_{\text{aver}} = -0.08$ ,  $SD_{\text{aver}} = 0.21$ ,  $M_{\text{non-aver}} = -0.06$ ,  $SD_{\text{non-aver}} = 0.21$ ) modality (Table 1, Fig. 3). After multiple comparisons correction, HR responses to aversive versus non-aversive stimuli showed a significant differentiation from stimulus onset to the end of the examined interval (0-7s) for

audiovisual stimuli, and from 1s after stimulus onset to stimulus offset (1-5s) for auditory stimuli.

### *Electromyography*

*Corrugator supercilii.* We examined differences in EMG<sub>C</sub> for aversive and non-aversive conditions across the three modalities. There was an interaction effect of aversiveness and modality (Table 1). Planned comparisons were conducted to examine the aversiveness effect for each modality. Consistent with an enhancement effect, EMG<sub>C</sub> potentiation was significantly greater for the aversive compared to non-aversive conditions for audiovisual ( $M_{\text{aver}} = 0.10$ ,  $SD_{\text{aver}} = 0.27$ ,  $M_{\text{non-aver}} = -0.06$ ,  $SD_{\text{non-aver}} = 0.27$ ) and auditory ( $M_{\text{aver}} = 0.02$ ,  $SD_{\text{aver}} = 0.28$ ,  $M_{\text{non-aver}} = -0.11$ ,  $SD_{\text{non-aver}} = 0.26$ ), but not for visual ( $M_{\text{aver}} = 0.03$ ,  $SD_{\text{aver}} = 0.23$ ,  $M_{\text{non-aver}} = -0.002$ ,  $SD_{\text{non-aver}} = 0.24$ ) modality (Table 1, Fig.3). After multiple comparisons correction, EMG<sub>C</sub> responses to aversive versus non-aversive stimuli showed a significant differentiation from 1s after the stimulus onset to the end of the examined interval (1-7s) for both audiovisual and auditory stimuli.

*Zygomaticus major.* We examined differences in EMG<sub>Z</sub> for aversive and non-aversive conditions across the three modalities. There were main effects of aversiveness and modality (Table 1). Planned comparisons were conducted to examine the aversiveness effect for each modality. EMG<sub>Z</sub> potentiation was significantly greater for the aversive compared to non-aversive conditions for audiovisual ( $M_{\text{aver}} = 0.07$ ,  $SD_{\text{aver}} = 0.21$ ,  $M_{\text{non-aver}} = -0.01$ ,  $SD_{\text{non-aver}} = 0.20$ ), but not for auditory ( $M_{\text{aver}} = 0.08$ ,  $SD_{\text{aver}} = 0.22$ ,  $M_{\text{non-aver}} = 0.01$ ,  $SD_{\text{non-aver}} = 0.21$ ) or visual ( $M_{\text{aver}} = -0.07$ ,  $SD_{\text{aver}} = 0.18$ ,  $M_{\text{non-aver}} = -0.10$ ,  $SD_{\text{non-aver}} = 0.20$ ) modalities (Table 1, Fig. 3). After multiple comparisons correction, EMG<sub>Z</sub> responses to aversive versus non-aversive stimuli showed a

significant differentiation from the stimulus offset to the end of the examined interval (5-7s) for audiovisual stimuli.

### *Electrodermal Activity*

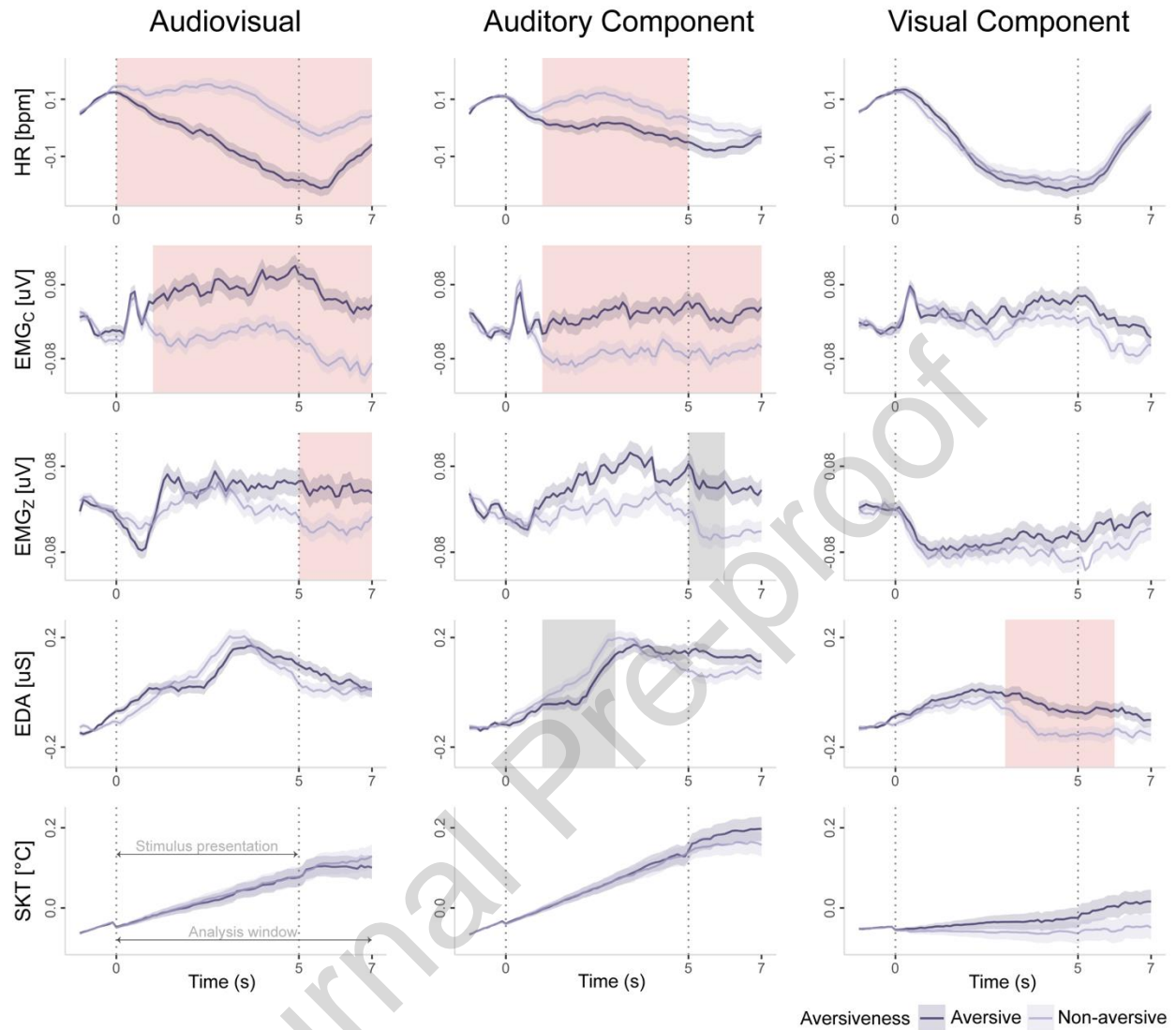
We examined differences in electrodermal activity for aversive and non-aversive conditions across the three modalities. Only the main effect of modality was significant (Table 1). Planned comparisons were conducted to examine aversiveness effect for each modality. Phasic EDA was significantly greater for the aversive compared to non-aversive conditions for visual ( $M_{\text{aver}} = -0.13$ ,  $SD_{\text{aver}} = 0.20$ ,  $M_{\text{non-aver}} = -0.20$ ,  $SD_{\text{non-aver}} = 0.19$ ), but not for audiovisual ( $M_{\text{aver}} = 0.04$ ,  $SD_{\text{aver}} = 0.18$ ,  $M_{\text{non-aver}} = 0.01$ ,  $SD_{\text{non-aver}} = 0.28$ ) or auditory ( $M_{\text{aver}} = 0.06$ ,  $SD_{\text{aver}} = 0.21$ ,  $M_{\text{non-aver}} = 0.05$ ,  $SD_{\text{non-aver}} = 0.20$ ) modalities (Table 1, Fig. 3). After multiple comparisons correction, EDA responses to aversive versus non-aversive stimuli showed a significant differentiation from 3s to 6s after the stimulus onset for visual stimuli.

### *Skin Temperature*

We examined skin temperature differences between aversive and non-aversive conditions across the three modalities. The only significant effect was the main effect of modality (Table 1). Planned comparisons revealed no significant results for differences between aversive and non-aversive conditions (Table 1, Fig. 3).

## **Figure 3**

### *Physiological measures across all conditions*



**Notes. A.** Visualization of the z-transformed signal changes across time in the analysis window. Solid lines for aversive (darker) and non-aversive (lighter) conditions indicate signals averaged across trials and participants with corresponding 95% confidence intervals. Dotted vertical lines indicate stimulus onset and offset. The analysis time window was 0-7s after the stimulus onset. Time windows for statistically significant differences between aversive and non-aversive conditions that survived the permutation test are shaded in pink. Time windows for identified clusters that did not survive the permutation test are shaded in grey.

### Residualized analyses controlling for arousal

We ran the same ANOVA and planned comparisons using the residuals of physiological signals after removing the variance explained by arousal ratings. The pattern of significance for heart rate remained unchanged for both the overall F-test and planned comparisons. For EMG<sub>C</sub>, a significant modality effect emerged after controlling for arousal. However, the pattern of results from planned comparisons remained unchanged. For EMG<sub>Z</sub>, the planned comparison of aversiveness for audiovisual modality was no longer significant after controlling for arousal. For EDA, the pattern of significant overall F-test results remained unchanged, but the planned comparison result for the visual condition was no longer significant. For skin temperature, the pattern of results remained unchanged for both overall F-test and planned comparisons. Thus, HR and EMG<sub>C</sub> sensitivity to aversiveness could not be explained by arousal alone.

**Table 1. Statistical Table**

Figure	DV	Statistical Test	Statistical values	MSE / 95% CI	Significance	Effect Size
2	Valence	ANOVA				
		Aversiveness	$F_{(1,72)} = 1976.65$	$MSE = 0.88$	$p < .001^{***}$	$\eta_p^2 = .965$
		Modality	$F_{(2,144)} = 49.38$	$MSE = 0.22$	$p_{-GG} < .001^{***}$	$\eta_p^2 = .407$
		Interaction	$F_{(2,144)} = 89.24$	$MSE = 0.24$	$p < .001^{***}$	$\eta_p^2 = .553$
		Planned t-test				
		AV <sub>aver - non-aver</sub>	$t_{(72)} = -59.95$	[-4.99, -4.67]	$p < .001^{***}$	$d = 7.02$
		A <sub>aver - non-aver</sub>	$t_{(72)} = -29.62$	[-3.99, -3.49]	$p < .001^{***}$	$d = 3.47$
		V <sub>aver - non-aver</sub>	$t_{(72)} = -28.27$	[-3.63, -3.15]	$p < .001^{***}$	$d = 3.31$
2	Arousal	ANOVA				
		Aversiveness	$F_{(1,72)} = 17.82$	$MSE = 2.80$	$p < .001^{***}$	$\eta_p^2 = .198$
		Modality	$F_{(2,144)} = 130.17$	$MSE = 1.97$	$p_{-GG} < .001^{***}$	$\eta_p^2 = .644$
		Interaction	$F_{(2,144)} = 4.90$	$MSE = 0.28$	$p_{-GG} = .012^*$	$\eta_p^2 = .064$
		Planned t-test				
		AV <sub>aver - non-aver</sub>	$t_{(72)} = 3.42$	[0.30, 1.15]	$p = .001^{**}$	$d = 0.40$

		$A_{\text{aver - non-aver}}$	$t_{(72)} = 5.26$	[0.51, 1.14]	$p < .001^{***}$	$d = 0.62$
		$V_{\text{aver - non-aver}}$	$t_{(72)} = 3.32$	[0.19, 0.76]	$p < .01^{**}$	$d = 0.39$
3	HR	ANOVA				
		Aversiveness	$F_{(1,71)} = 23.60$	$MSE = 0.05$	$p < .001^{***}$	$\eta_p^2 = .249$
		Modality	$F_{(2,142)} = 15.23$	$MSE = 0.06$	$p < .001^{***}$	$\eta_p^2 = .177$
		Interaction	$F_{(2,142)} = 5.33$	$MSE = 0.05$	$p = .006^{**}$	$\eta_p^2 = .070$
		<i>Planned t-test</i>				
		$AV_{\text{aver - non-aver}}$	$t_{(71)} = -5.16$	[-0.27, -0.12]	$p < .001^{***}$	$d = 0.61$
		$A_{\text{aver - non-aver}}$	$t_{(71)} = -2.45$	[-0.17, -0.02]	$p < .017^*$	$d = 0.29$
		$V_{\text{aver - non-aver}}$	$t_{(71)} = -0.51$	[-0.10, 0.06]	$p = .611$	-
3	EMG <sub>C</sub>	ANOVA				
		Aversiveness	$F_{(1,71)} = 10.03$	$MSE = 0.13$	$p = .002^{**}$	$\eta_p^2 = .124$
		Modality	$F_{(2,142)} = 2.30$	$MSE = 0.08$	$p = .105$	-
		Interaction	$F_{(2,142)} = 3.94$	$MSE = 0.04$	$p = .022^*$	$\eta_p^2 = .053$
		<i>Planned t-test</i>				
		$AV_{\text{aver - non-aver}}$	$t_{(71)} = 3.31$	[0.06, 0.26]	$p = .002^*$	$d = 0.39$
		$A_{\text{aver - non-aver}}$	$t_{(71)} = 2.72$	[0.04, 0.23]	$p = .008^*$	$d = 0.32$
		$V_{\text{aver - non-aver}}$	$t_{(71)} = 1.04$	[-0.03, 0.10]	$p = .303$	-
3	EMG <sub>Z</sub>	ANOVA				
		Aversiveness	$F_{(1,70)} = 5.21$	$MSE = 0.07$	$p = .025^*$	$\eta_p^2 = .069$
		Modality	$F_{(2,140)} = 18.91$	$MSE = 0.04$	$p < .001^{***}$	$\eta_p^2 = .213$
		Interaction	$F_{(2,140)} = 0.89$	$MSE = 0.03$	$p = .410$	-
		<i>Planned t-test</i>				
		$AV_{\text{aver - non-aver}}$	$t_{(70)} = 2.02$	[0.00, 0.15]	$p = .047^*$	$d = 0.24$
		$A_{\text{aver - non-aver}}$	$t_{(70)} = 1.79$	[-0.01, 0.15]	$p = .078$	-
		$V_{\text{aver - non-aver}}$	$t_{(70)} = 1.02$	[-0.02, 0.08]	$p = .313$	-
3	EDA	ANOVA				
		Aversiveness	$F_{(1,72)} = 2.03$	$MSE = 0.08$	$p = .159$	-
		Modality	$F_{(2,144)} = 39.93$	$MSE = 0.05$	$p < .001^{***}$	$\eta_p^2 = .357$
		Interaction	$F_{(2,144)} = 0.84$	$MSE = 0.04$	$p = .431$	-
		<i>Planned t-test</i>				
		$AV_{\text{aver - non-aver}}$	$t_{(72)} = 0.71$	[-0.06, 0.12]	$p = .479$	-
		$A_{\text{aver - non-aver}}$	$t_{(72)} = 0.28$	[-0.06, 0.08]	$p = .784$	-
		$V_{\text{aver - non-aver}}$	$t_{(72)} = 2.23$	[0.01, 0.13]	$p = .029^*$	$d = 0.26$
3	SKT	ANOVA				
		Aversiveness	$F_{(1,72)} = 0.03$	$MSE = 0.05$	$p = .859$	-
		Modality	$F_{(2,144)} = 10.63$	$MSE = 0.08$	$p_{-GG} < .001^{***}$	$\eta_p^2 = .129$
		Interaction	$F_{(2,144)} = 0.11$	$MSE = 0.05$	$p = .887$	-
		<i>Planned t-test</i>				

$AV_{\text{aver} - \text{non-aver}}$	$t_{(72)} = -0.21$	$[-0.08, 0.06]$	$p = .837$	-
$A_{\text{aver} - \text{non-aver}}$	$t_{(72)} = -0.43$	$[-0.08, 0.05]$	$p = .669$	-
$V_{\text{aver} - \text{non-aver}}$	$t_{(72)} = 0.24$	$[-0.07, 0.09]$	$p = .815$	-

*Note.* Two-way repeated measures ANOVAs were conducted for all dependent variables, along with planned *t*-tests. DV: dependent variable, aver: aversive condition, non-aver: non-aversive condition, AV: audiovisual, A: auditory, V: visual.  $p_{\text{GG}}$  indicates the  $p$  value after the Greenhouse-Geiser adjustment.  $P$ -values for planned  $t$ -tests were not adjusted. \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ .

## Discussion

Our study was designed to investigate the differences in physiological responses to aversiveness in naturalistic audiovisual stimuli and their auditory and visual components. Five physiological signals were acquired during naturalistic stimuli perception. Three of our measures (HR, EMG<sub>C</sub>, EMG<sub>Z</sub>) were sensitive to aversiveness in the audiovisual stimuli, with HR and EMG<sub>C</sub> also sensitive to the auditory component. Only EDA was sensitive to the visual component. SKT was not sensitive to the aversiveness manipulation for any modality. The sensitivity of HR and EMG<sub>C</sub> to aversiveness remained significant after controlling for arousal. The enhancement effect was observed in our ratings on valence and in the HR, EMG<sub>C</sub>, and EMG<sub>Z</sub> responses. Four of our measures (HR, EMG<sub>Z</sub>, EDA, SKT) were sensitive to modality, which may reflect valence and arousal differences across modalities. We discuss these findings below starting with the four hypotheses that we formulated.

In support of our first hypothesis, both EMG<sub>C</sub> and EMG<sub>Z</sub> showed greater potentiation to aversive audiovisual stimuli. (Bradley & Lang, 2000; Höfling et al., 2020; Larsen et al., 2003; Neta et al., 2009; Sato & Kochiyama, 2022) Our second hypothesis was partially supported. The

HR deceleration was larger for aversive compared to non-aversive audiovisual stimuli, but EDA potentiation was not. Our third hypothesis that aversiveness sensitivity would occur across modalities was partially supported. Valence and arousal ratings differed between aversive and non-aversive conditions across all three modalities. HR deceleration and EMG<sub>C</sub> potentiation were greater for aversive compared to non-aversive stimuli for the audiovisual and auditory modalities, but not for the visual modality. In contrast, EDA signal was greater for aversive compared to non-aversive stimuli for the visual modality only. EMG<sub>Z</sub> did not show greater potentiation to aversive compared to non-aversive stimuli in either unimodal condition. The fourth hypothesis that self-report and physiological sensitivity to aversiveness differences should be larger for the audiovisual modality was partially supported. Results from HR, EMG<sub>C</sub>, EMG<sub>Z</sub> and valence rating were consistent with an enhancement effect, i.e., amplification of the affective responses when unimodal information is integrated (Gao et al., 2018a, Gao et al., 2018b). However, the enhancement effect was not observed in other channels (EDA, SKT). (Bradley & Lang, 2000; Šolcová & Lačev, 2017; Salazar-López et al., 2015; Sato & Kochiyama, 2022)

The robust relationship between EMG<sub>C</sub> and aversiveness is in line with previous evidence that it responds stronger for negative valence (Bradley & Lang, 2000; Höfling et al., 2020; Larsen et al., 2003; Neta et al., 2009; Sato & Kochiyama, 2022). It may seem odd that EMG<sub>Z</sub> also showed greater potentiation for aversive stimuli. Literature about the relationship between zygomaticus major and valence is mixed. Some previous findings suggest that zygomaticus major muscle is an indicator of pleasantness (Baur et al., 2015; Lang et al., 1993, Sato et al., 2008). Other findings suggest there is a quadratic relationship between zygomaticus major potentiation and valence (Kappas & Pecchinenda, 1998; Larsen et al., 2003; Menne & Schwab, 2018). One possible interpretation is that zygomaticus major potentiation is tied to avoidance



behaviors, like disgust (Shook et al., 2019). Previous findings have found heightened EMGz potentiation when processing disgusting pictures (Lang et al., 1993) and dynamic videos (De Jong et al., 2002). Another possible interpretation is that zygomatic muscles control multiple mouth movement angles that allow for diverse facial expressions, including lip corner puller (Action Unit 12; zygomaticus major) and nasolabial deepener (Action Unit 11; zygomaticus minor), the two action units related to hedonic valence. Tan et al. (2012) pointed out the possibility of the influence of cross talk from multiple muscles adjacent to the zygomaticus major. A future study can include measurement of both zygomaticus major and minor to better understand their roles in responding to aversiveness.

We found heightened HR deceleration following the stimulus presentation to aversive versus non-aversive conditions for both the audiovisual stimuli and their auditory components. These results are in line with previous findings that initial heart rate deceleration was related to aversiveness perception across various contexts (Bradley et al., 2001; Bradley & Lang, 2000; Gomez et al., 2016; Lang et al., 1993; Palomba et al., 2000). Surprisingly, the significant HR difference between aversive and non-aversive stimuli found for audiovisual and audio presentations was not absent in the visual modality. As shown in the line graphs (Figure 3), there was a sharp HR deceleration for aversive visual stimuli that was as strong or stronger than that for the aversive audiovisual and auditory presentations. However, non-aversive visual stimuli also showed a prominent deceleration (0-5s), rather than the slight initial deceleration (0-1s), subsequent acceleration (1-3s), and later deceleration (3-5s) pattern found for non-aversive audiovisual and auditory stimuli. The HR acceleration phase occurring 2-3s after stimuli presentation has been shown to be a characteristic feature of both pleasant sounds (Bradley & Lang, 2000) and pictures (Bradley et al., 2001; Lang et al., 1993). Thus, one might argue that the

lack of significant HR difference for the visual component in our data may reflect the presence of sustained HR deceleration rather than the expected initial HR deceleration and subsequent acceleration. Bradley et al. (2001) suggested that the HR acceleration in response to appetitive stimuli may indicate the “success of a recognition-encoding process”, with more readily identifiable stimuli (e.g., food and nature) eliciting greater acceleration compared to stimuli requiring more sustained processing (e.g., families or sports scenes with novel individuals). A recent review (Skora et al., 2022) proposes that HR deceleration is associated with enhanced attention to external stimuli, while HR acceleration is associated with enhanced action tendencies and behavioral readjustment. One potential explanation for the strong HR deceleration observed for non-aversive visual stimuli in our study is that identifying videos without sound may generally demand greater attentional resources compared to audiovisual clips and sounds alone.

Our results on EDA only showed that it was sensitive to aversiveness for the visual component. This may suggest an effect of arousal, instead of valence. There are many studies that has shown a reliable relationship between EDA and arousal (e.g., Lang et al., 1993; Sato et al., 2020; Sato & Kochiyama, 2022). In partial support of this interpretation, our EDA results were no longer significant once arousal was accounted for.

Because we focused on aversiveness in naturalistic stimuli, valence and arousal were inherently confounded (Kuppens et al., 2013), unlike controlled stimuli that are specifically designed to be balanced in terms of both valence and arousal. This natural overlap between valence and arousal in real-world stimuli makes it challenging to isolate the individual contributions of each factor with stimulus manipulation. To better understand our significant results of aversiveness sensitivity, we regressed out the effect of arousal to statistically isolate the effect of valence using residualized ANOVA. Our results revealed that the sensitivity of heart

rate and corrugator supercilii to aversiveness remain even when controlling for arousal. This suggests that valence is the primary contributor to the physiological processes observed in our study.

Many real-life scenarios involve sounds that carry emotional weight, such as the sounds of honking horns on a busy street. Our stimuli emphasize sensory sensitivity to sound, while visual components tend to be less emotionally evocative. For example, it is the jarring noise of a chair scraping across the floor that creates an aversive reaction, rather than the visual of the scene itself. Thus, it is unsurprising for our stimuli set that the aversiveness was detectable in heart rate and corrugator supercilii signals for auditory component, and not for the visual component. These results highlight the powerful role that sound plays in shaping emotional experiences.

Our study extended previous research regarding the effects of presentation modality by comparing naturalistic audiovisual clips to their isolated unimodal components. This approach allowed us to examine the enhancement effect across modalities more directly while controlling for content differences. Our self-report data replicated the enhancement effect, in which the difference in valence for aversive versus non-aversive stimuli was greater for the combined audiovisual stimuli than for either modality separately (Baumgartner et al., 2006; Gao, Wedell, Green, et al., 2018; Gao, Wedell, Kim, et al., 2018). The observed interaction of aversiveness and modality for HR and EMG<sub>C</sub> also supported this effect in physiological data.

There are a number of limitations in the current study. Our participants were predominantly female college students, reflecting student composition of the psychology participant pool. Thus, we were not able to examine the effects of age and gender (Burris et al., 2007) on physiological responses to aversiveness. Also, we used a fairly short stimuli

presentation interval (5s). However, past research in our lab has identified valence and arousal from fMRI (Kim et al., 2016, 2017) and physiological (Kim & Wedell, 2016) signals using 4s, 5s, and 8s presentation intervals. Although the naturalistic stimuli, like the ones used in this study, have the advantage of increased ecological validity, they lack control over lower-level auditory and visual features that may contribute to physiological responses (Coutinho & Cangelosi, 2011; Delplanque et al., 2007). In addition, the direct comparison with prior research is limited given no extreme content (e.g. mutilation, erotica themes) or food related themes that previously elicited strong cardiac responses (Bradley et al., 2001) were used in the present study. Finally, by comparing aversive and non-aversive stimuli, we are adopting a bipolar representation of valence and cannot test for effects that may emerge from bivalent approach. Further studies could address some of these limitations.

In conclusion, the present study isolated the unimodal auditory and visual components from naturalistic audiovisual clips to examine physiological responses as a function of aversiveness and presentation modality. We found that heart rate and corrugator supercilii potentiation were consistently sensitive to aversiveness across audiovisual and auditory modalities. Electrodermal activity and zygomaticus major signals were less reliably related to aversiveness. Overall, physiological signals for the visual components were less differentiated. In addition, our results showed partial support for an emotional enhancement effect for cross-modal integration in both physiological activity and affective ratings. These findings extended previous research on psychophysiological audiovisual pairing and provide insights for future studies that may sample from clinical populations.

### **Funding**

This work was supported by a grant from the Misophonia Research Fund.

### **Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. There are no patents or additional relationships to disclose.

Declaration of Generative AI and AI-assisted technologies in the writing process

Statement: The authors did not use generative AI technologies for preparation of this work.

## References

- Amrhein, C., Mühlberger, A., Pauli, P., & Wiedemann, G. (2004). Modulation of event-related brain potentials during affective picture processing: A complement to startle reflex and skin conductance response? *International Journal of Psychophysiology*, 54(3), 231–240.  
<https://doi.org/10.1016/j.ijpsycho.2004.05.009>
- Baldaro, B., Tuoizzi, G., Codispoti, M., Montebanocci, O., Barbagli, F., Trombini, E., & Rossi, N. (2004). Aggressive and non-violent videogames: Short-term psychological and cardiovascular effects on habitual players. *Stress and Health*, 20(4), 203–208.  
<https://doi.org/10.1002/smi.1015>
- Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, 60(1), 34–43. <https://doi.org/10.1016/j.ijpsycho.2005.04.007>
- Baur, R., Conzelmann, A., Wieser, M. J., & Pauli, P. (2015). Spontaneous emotion regulation: Differential effects on evoked brain potentials and facial muscle activity. *International Journal of Psychophysiology*, 96(1), 38–48. <https://doi.org/10.1016/j.ijpsycho.2015.02.022>
- Bernat, E., Patrick, C. J., Benning, S. D., & Tellegen, A. (2006). Effects of picture content and intensity on affective physiological response. *Psychophysiology*, 43(1), 93–103.  
<https://doi.org/10.1111/j.1469-8986.2006.00380.x>
- Bradley, M. M., & Lang, P. J. (2006). Motivation and emotion. In J.T. Cacioppo, L. G. Tassinary, and G. Berntson (Eds.). *Handbook of Psychophysiology* (2<sup>nd</sup> Edition), 581–607. New York: Cambridge University Press.
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and Motivation I:

- Defensive and Appetitive Reactions in Picture Processing. *Emotion*, 1(3), 276–298.  
<https://doi.org/10.1037/1528-3542.1.3.276>
- Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2), 204–215. <https://doi.org/10.1017/S0048577200990012>
- Burris, L., Powell, D. A., & White, J. (2007). Psychophysiological and subjective indices of emotion as a function of age and gender. *Cognition and Emotion*, 21(1), 182–210.  
<https://doi.org/10.1080/02699930600562235>
- Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic Activity Over Facial Muscle Regions Can Differentiate the Valence and Intensity of Affective Reactions. *Journal of Personality and Social Psychology*, 50(2), 260–268.  
<https://doi.org/10.1037/0022-3514.50.2.260>
- Carver, C. S., & Harmon-Jones, E. (2009). Anger Is an Approach-Related Affect: Evidence and Implications. *Psychological Bulletin*, 135(2), 183–204. <https://doi.org/10.1037/a0013965>
- Codispoti, M., Surcinelli, P., & Baldaro, B. (2008). Watching emotional movies: Affective reactions and gender differences. *International Journal of Psychophysiology*, 69(2), 90–95.  
<https://doi.org/10.1016/j.ijpsycho.2008.03.004>
- Coutinho, E., & Cangelosi, A. (2011). Musical Emotions: Predicting Second-by-Second Subjective Feelings of Emotion From Low-Level Psychoacoustic Features and Physiological Measurements. *Emotion*, 11(4), 921–937. <https://doi.org/10.1037/a0024700>
- Crone, E. A., Somsen, R. J. M., Van Beek, B., & Van Der Molen, M. W. (2004). Heart rate and skin conductance analysis of antecedents and consequences of decision making. *Psychophysiology*, 41(4), 531–540. <https://doi.org/10.1111/j.1469-8986.2004.00197.x>
- De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological

validity. *Trends in Cognitive Sciences*, 7(10), 460–467.

<https://doi.org/10.1016/j.tics.2003.08.014>

De Jong, P. J., Peters, M., & Vanderhallen, I. (2002). Disgust and disgust sensitivity in spider phobia: Facial EMG in response to spider and oral disgust imagery. *Journal of Anxiety Disorders*, 16(5), 477–493. [https://doi.org/10.1016/S0887-6185\(02\)00167-6](https://doi.org/10.1016/S0887-6185(02)00167-6)

Dellacherie, D., Roy, M., Hugueville, L., Peretz, I., & Samson, S. (2011). The effect of musical experience on emotional self-reports and psychophysiological responses to dissonance. *Psychophysiology*, 48(3), 337–349. <https://doi.org/10.1111/j.1469-8986.2010.01075.x>

Edelstein, M., Brang, D., Rouw, R., & Ramachandran, V. S. (2013). Misophonia: Physiological investigations and case descriptions. *Frontiers in Human Neuroscience*, 7, 1–11. <https://doi.org/10.3389/fnhum.2013.00296>

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

Gao, C., Wedell, D. H., Green, J. J., Jia, X., Mao, X., Guo, C., & Shinkareva, S. V. (2018). Temporal dynamics of audiovisual affective processing. *Biological Psychology*, 139(August), 59–72. <https://doi.org/10.1016/j.biopsycho.2018.10.001>

Gao, C., Wedell, D. H., Kim, J., Weber, C. E., & Shinkareva, S. V. (2018). Modelling audiovisual integration of affect from videos and music. *Cognition and Emotion*, 32(3), 516–529. <https://doi.org/10.1080/02699931.2017.1320979>

Garrett, A. S., & Maddock, R. J. (2001). Time course of the subjective emotional response to aversive pictures: Relevance to fMRI studies. *Psychiatry Research - Neuroimaging*, 108(1), 39–48. [https://doi.org/10.1016/S0925-4927\(01\)00110-X](https://doi.org/10.1016/S0925-4927(01)00110-X)



- Gjoreski, M., Kiprijanovska, I., Stankoski, S., Mavridou, I., Broulidakis, M. J., Gjoreski, H., & Nduka, C. (2022). Facial EMG sensing for monitoring affect using a wearable device. *Scientific Reports*, 12(1), 1–12. <https://doi.org/10.1038/s41598-022-21456-1>
- Gomez, P., von Gunten, A., & Danuser, B. (2016). Autonomic nervous system reactivity within the valence–arousal affective space: Modulation by sex and age. *International Journal of Psychophysiology*, 109, 51–62. <https://doi.org/10.1016/j.ijpsycho.2016.10.002>
- Greco, A., Valenza, G., Citi, L., & Scilingo, E. P. (2017). Arousal and valence recognition of affective sounds based on electrodermal activity. *IEEE Sensors Journal*, 17(3), 716–725. <https://doi.org/10.1109/JSEN.2016.2623677>
- Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgement and psychophysiological response: dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology*, 3, 51–64.
- Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Emotions Over Time: Synchronicity and Development of Subjective, Physiological, and Facial Affective Reactions to Music. *Emotion*, 7(4), 774–788. <https://doi.org/10.1037/1528-3542.7.4.774>
- Hayes, W. M., & Wedell, D. H. (2020). Autonomic responses to choice outcomes: Links to task performance and reinforcement-learning parameters. *Biological Psychology*, 156, 107968. <https://doi.org/10.1016/j.biopsycho.2020.107968>
- Höfling, T. T. A., Gerdes, A. B. M., Föhl, U., & Alpers, G. W. (2020). Read My Face: Automatic Facial Coding Versus Psychophysiological Indicators of Emotional Valence and Arousal. *Frontiers in Psychology*, 11, 1–15. <https://doi.org/10.3389/fpsyg.2020.01388>
- Jäncke, L., Vogt, J., Musial, F., Lutz, K., & Kalveram, K. T. (1996). Facial EMG responses to auditory stimuli. *International Journal of Psychophysiology*, 22, 85–96.

[https://doi.org/10.1016/0167-8760\(96\)00013-X](https://doi.org/10.1016/0167-8760(96)00013-X)

- Jung, J., Zbozinek, T. D., Cummings, K. K., Wilhelm, F. H., Dapretto, M., Craske, M. G., Bookheimer, S. Y., & Green, S. A. (2021). Associations between physiological and neural measures of sensory reactivity in youth with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 62(10), 1183–1194. <https://doi.org/10.1111/jcpp.13387>
- Kappas, A., & Pecchinenda, A. (1998). Zygomaticus Major activity is not selective indicator of positive affective state in ongoing interactive tasks. *Psychophysiology*, 35.
- Kim, J., Shinkareva, S. V., & Wedell, D. H. (2017). Representations of modality-general valence for videos and music derived from fMRI data. *NeuroImage*, 148, 42–54. <https://doi.org/10.1016/j.neuroimage.2017.01.002>
- Kim, J., Wang, J., Wedell, D. H., & Shinkareva, S. V. (2016). Identifying core affect in individuals from fMRI responses to dynamic naturalistic audiovisual stimuli. *PLoS ONE*, 11(9), 1–21. <https://doi.org/10.1371/journal.pone.0161589>
- Kim, J., & Wedell, D. H. (2016). Comparison of physiological responses to affect eliciting pictures and music. *International Journal of Psychophysiology*, 101, 9–17. <https://doi.org/10.1016/j.ijpsycho.2015.12.011>
- Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, 139(4), 917–940. <https://doi.org/10.1037/a0030811>
- Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, 84(3), 437–450. <https://doi.org/10.1016/j.biopsycho.2009.10.007>
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261–273.

<https://doi.org/10.1111/j.1469-8986.1993.tb03352.x>

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii.

*Psychophysiology*, 40(5), 776–785. <https://doi.org/10.1111/1469-8986.00078>

Lazar, J. N., & Pearlman-Avnion, S. (2014). Effect of Affect Induction Method on Emotional Valence and Arousal. *Psychology*, 5, 595–601. <https://doi.org/10.4236/psych.2014.57070>

Menne, I. M., & Schwab, F. (2018). Faces of Emotion: Investigating Emotional Facial Expressions Towards a Robot. *International Journal of Social Robotics*, 10(2), 199–209.

<https://doi.org/10.1007/s12369-017-0447-2>

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator Muscle Responses Are Associated with Individual Differences in Positivity-Negativity Bias. *Emotion*, 9(5), 640–648.

<https://doi.org/10.1037/a0016819.Corrugator>

Nitschke, J. B., Sarinopoulos, I., MacKiewicz, K. L., Schaefer, H. S., & Davidson, R. J. (2006). Functional neuroanatomy of aversion and its anticipation. *NeuroImage*, 29(1), 106–116.

<https://doi.org/10.1016/j.neuroimage.2005.06.068>

Palomba, D., Sarlo, M., Angrilli, A., Mini, A., & Stegagno, L. (2000). Cardiac responses associated with affective processing of unpleasant film stimuli. *International Journal of*

*Psychophysiology*, 36(1), 45–57. [https://doi.org/10.1016/S0167-8760\(99\)00099-9](https://doi.org/10.1016/S0167-8760(99)00099-9)

Rosa, P. J., Oliveira, J., Alghazzawi, D., Fardoun, H., & Gamito, P. (2017). Affective and physiological correlates of the perception of unimodal and bimodal emotional stimuli.

*Psicothema*, 29(3), 364–369. <https://doi.org/10.7334/psicothema2016.272>

Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 145–172. <https://doi.org/10.1037/0033-295X.110.1.145>

- Salazar-López, E., Domínguez, E., Juárez Ramos, V., de la Fuente, J., Meins, A., Iborra, O., Gálvez, G., Rodríguez-Artacho, M. A., & Gómez-Milán, E. (2015). The mental and subjective skin: Emotion, empathy, feelings and thermography. *Consciousness and Cognition*, 34, 149–162. <https://doi.org/10.1016/j.concog.2015.04.003>
- Sato, W., & Kochiyama, T. (2022). Exploration of Emotion Dynamics Sensing Using Trapezius EMG and Fingertip Temperature. *Sensors*, 22(17), 1–15. <https://doi.org/10.3390/s22176553>
- Sato, W., Kochiyama, T., & Yoshikawa, S. (2020). Physiological correlates of subjective emotional valence and arousal dynamics while viewing films. *Biological Psychology*, 157, 107974. <https://doi.org/10.1016/j.biopsycho.2020.107974>
- Scott Killgore, W. D. (1998). The affect grid: A moderately valid, nonspecific measure of pleasure and arousal. *Psychological Reports*, 83(2), 639–642. <https://doi.org/10.2466/pr0.1998.83.2.639>
- Shook, N. J., Thomas, R., & Ford, C. G. (2019). Testing the relation between disgust and general avoidance behavior. *Personality and Individual Differences*, 150, 149–163. <https://doi.org/10.1016/j.paid.2019.05.063>
- Singh, T., & Seidman, M. D. (2019). Hearing disorders associated with mild traumatic brain injury (mTBI). In *Neurosensory Disorders in Mild Traumatic Brain Injury*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-812344-7.00011-X>
- Skora, L. I., Livermore, J., J., A., & Roelofs, K. (2022). The functional role of cardiac activity in perception and action. *Neuroscience and Biobehavioral Reviews*, 137, 104655. <https://doi.org/10.1016/j.neubiorev.2022.104655>
- Šolcová, P. I., & Lačev, A. (2017). Differences in male and female subjective experience and physiological reactions to emotional stimuli. *International Journal of Psychophysiology*,

- 117, 75–82. <https://doi.org/10.1016/j.ijpsycho.2017.04.009>
- Swedo, S. E., Baguley, D. M., Denys, D., Dixon, L. J., Erfanian, M., Fioretti, A., Jastreboff, P. J., Kumar, S., Rosenthal, M. Z., Rouw, R., Schiller, D., Simner, J., Storch, E. A., Taylor, S., Werff, K. R. V., Altimus, C. M., & Raver, S. M. (2022). Consensus Definition of Misophonia: A Delphi Study. *Frontiers in Neuroscience*, 16, 1–16. <https://doi.org/10.3389/fnins.2022.841816>
- Tan, J. W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *Journal of Ambient Intelligence and Humanized Computing*, 3(1), 3–10. <https://doi.org/10.1007/s12652-011-0084-9>
- Vowles, K. E., McNeil, D. W., Sorrell, J. T., & Lawrence, S. M. (2006). Fear and pain: Investigating the interaction between aversive states. *Journal of Abnormal Psychology*, 115(4), 821–833. <https://doi.org/10.1037/0021-843X.115.4.821>
- Tan, J. W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *J Ambient Intell Human Comput*, 3, 3–10. <https://doi.org/10.1007/s12652-011-0084-9>
- Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record*, 58(3), 475–482. <https://doi.org/10.1007/BF03395630>

## **Declaration of Generative AI and AI-assisted technologies in the writing process**

No AI-assistant tools were used in the preparation of this article.

### Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Svetlana Shinkareva reports financial support was provided by Misophonia Research Fund. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Highlights

Audiovisual clips and their components were examined in the same experiment

- Heart rate deceleration for aversive audiovisual clips and their auditory components
- Increased corrugator supercilii potentiation for aversive audiovisual clips
- Increased corrugator supercilii potentiation for aversive auditory components
- Overall, physiological responses to the visual component were less distinct