The impact of disgusting sounds on pupil diameter of misophonic and non-misophonic listeners

Urszula Oszczapinska[1], Sungjoon Park[1], Yuqi Qiu[1], Bridget Nance[1], Megan Julien[1], and Laurie Heller[1]

[1]Carnegie Mellon University

October 19, 2024

Abstract

jabbrv-ltwa-all.ldf jabbrv-ltwa-en.ldf Everyday sounds can elicit a range of emotional and physiological responses. For individuals with misophonia, some sounds can produce strong feelings of disgust, annoyance, and anger, often accompanied by increased perspiration and heart rate. Presently, methods of diagnosing misophonia rely on clinical interviews and self-assessment scales. Our study asks whether pupillometry can be an objective measure that correlates with self-reported misophonia severity. Previous studies show that both unpleasant and pleasant sounds increase pupil diameter (Partala and Surakka, 2003; Nakakoga et al., 2020); however, these have not compared pupil responses to disgust versus other emotions. Given prior indications that the response to visually disgusting stimuli is pupil constriction (Ayzenberg et al., 2018), we asked whether the pupil dilation to auditorily disgusting stimuli would be smaller than for other emotional sounds. In our listening task, we monitored pupil size changes while participants listened to positive and negative emotional sounds from the IADS database (Bradley and Lang, 2007) along with "triggers" known to be especially aversive to misophonics. Participants reported the intensity of their emotional reactions (disgust, anger, annoyance, happiness, sadness, fear) as well as valence and arousal. Misophonic listeners reported greater emotion intensity for emotions associated with triggers (disgust, anger, annoyance) as well as for fear. For all listeners, there was a positive association between changes in pupil diameter and emotion intensity. Overall, misophonics had greater pupil dilation than non-misophonics, but after equating for emotion category of the sounds, misophonic pupil dilation was only larger for trigger sounds (and marginally, disgusting sounds).

I. Introduction

Everyday sounds serve as informative signals about the physical events, objects, and agents in our environment. The acoustic properties of sound, spectral content, temporal content, and sound level are physical correlates of perceptual features, providing listeners with the ability to discern the size and material (Houben, 2002; Grassi, 2005; Klatzky, Pai and Krotkov, 2000; Freed, 1990) of various objects, as well as the action(s) imposed on them (Warren and Verbrugge, 1984; Lemaitre and Heller, 2012). Such properties are further influenced by our experiences, and expectations of where or when the sound is encountered. For example, listeners may interpret a set of repetitive sounds produced by wheels contacting a track as a train rolling. However, if an irregularity is present within the repetitive pattern of the sound, the listener may interpret the event as being produced by an object other than a train (Ballas and Howard, 1987). Distinguishing the temporal regularity of events is only one example of how acoustic features can inform sound identification. There are many other similar distinctions that listeners make day-to-day which ultimately influences what action should be made in response to a sound event.

It is also the case that everyday sounds can be a medium for expressing and inducing affective states. In music, changes in features such as the tempo, mode, timbre, and loudness have been observed to be the musical correlates for different emotions (Hunter and Schellenberg, 2010; Hevner, 1935; Ilie and Thompson, 2006). For example, 'happy-sounding' musical excerpts most often have a fast tempo and are played in a major mode (Gabrielsson and Juslin, 2003; Vieillard et al. 2008). In contrast, 'sad-sounding' musical excerpts most often have a slow tempo and are played in a minor mode, while 'fearful-sounding' excerpts contain an additional component of rhythmic irregularity and dissonance (Vieillard et al. 2008). In any of these cases, a heightened, or intense arousal response to music may result in a physiological reaction referred to as "musical chills" (Goldstein, 1980; Craig, 2005; Panksepp, 1995). Moreover, there is evidence that highly arousing music causes changes in heart rate (Sammler et al., 2007), and pupil diameter (Gringas et al., 2015) as well as changes in movement of facial cheek and brow muscles (Witvliet and Vrana, 2007; Khalfa et al., 2008; Lundqvist et al., 2009).

The link between sound properties and emotion in music holds true for general, everyday sounds as well. As an example, the sound of a sharp object scraping across a surface has been consistently reported as unpleasant; even the mere thought of nails scratching a blackboard causes listeners to wince or cringe (Halpern, Blake and Hillenbrand, 1986; Cox, 2008). Acoustic analysis of these sounds has indicated that mid-range frequencies (peak around 2 kHz) are the primary reason for why individuals find them unpleasant. In contrast, everyday sounds that are perceived as pleasant (e.g., water splashing) contain acoustic properties associated with timbre quality rich in softness and roundness (Ballas, 1993; Radsten Ekman, Lunden, and Nilsson, 2015). Physiologically, both unpleasant and pleasant environmental sounds alike reliably increase pupil diameter (Partala and Surakka, 2003; Nakakoga et al., 2020), heart rate (Nardelli et al., 2015; Bradley and Lang, 2000), perspiration (measured through skin conductance) (Bradley and Lang, 2000), and facial muscle movement (Bradley and Lang, 2000). In other words, sounds that differ in valence, but that are matched on arousal, elicit similar physiological reactions. However, there is also evidence that valence does contribute to the overall magnitude of a response (see Partala and Surakka, 2003). Thus, physiological reactions appear to scale in accordance with the perceived valence of the sound and the degree of arousal evoked.

Negative emotions elicited by certain sound events are not limited to acoustic causes; sometimes they are driven by the semantic representation of the event. Orofacial sounds like wet chewing, sniffing and heavy breathing, as well as non-orofacial, repetitive sounds like a clock-ticking or keyboard typing, have been reported to evoke high levels of disgust and anger for a specific population of individuals (Hansen et al., 2021; Swedo et al., 2022). These same individuals may experience negative physiological reactions when encountering these sounds (Jastreboff and Jastreboff, 2001. When reactions to any of these sounds produce unbearable levels of discomfort (Edelstein et al., 2013, Schroder et al., 2013), they can be symptoms of misophonia (Swedo et al., 2022). Note, individuals that do not have misophonia may still perceive these sounds to be unpleasant and disgusting, but not to the same degree of intensity. Thus, if such sounds show reliable physiological differences between the two populations, this category of sounds may be a good candidate for objectively diagnosing misophonia severity or status. Reported prevalence rates from studies using misophonic populations suggest that upward of 20% of the population may have some degree of these types of experiences to a set of trigger sounds. Unlike the general classes of sounds reviewed above, the sounds that evoke negative reactions do not have a strong acoustic basis (Kazazis, Korsmit, and McAdams, 2024) and are more strongly associated with the source event causing the sound (e.g. Heller and Smith, 2022). Promising research has suggested that reactions to these sounds are better explained by a social-cognitive mechanism. Given that most of these trigger sounds are produced by everyday actions by other individuals, it has been proposed that distress occurs as a result of not being able to prevent these sound-producing events from occurring (Ash et al., 2024). Others have proposed that misophonia is a result of a conditioned reflex (Jastreboff and Jastreboff, 2023).

Across the empirical studies assessing misophonic reactions, self-report scales and psychological evaluations have been relied on by researchers. These scales measure the intensity of an emotional response to a sound, such as disgust, anger, or anxiety. While misophonia research tends to contrast the aversiveness of triggers with other unpleasant sounds (see Hansen et al., 2021; Savard et al., 2022), no prior study has specifically examined the unique mix of underlying emotions experienced by misophonics in response to triggers and other disgusting sounds. Until recently, only a handful of studies have directly measured physiological responses, such as increases in heart rate and perspiration, in response to misophonic trigger sounds (e.g., Grossini et al., 2022; Siepsiak et al., 2023). However, no study to date investigating misophonic populations has utilized pupillometry as an alternative method to assessing responses to everyday sounds, in conjunction with emotion intensity ratings.

We attempt to bridge this gap by investigating both emotional and physiological responses to a range of negative and positive everyday sounds between two populations of young adults: individuals who scored at the clinical levels on misophonia severity scales and non-misophonic controls. In our study, we collected self-reported emotion intensity ratings (disgust, anger, annoyance, happiness, sadness, fear) as well as valence and arousal to a diverse set of sounds. By examining a range of self-reported emotions, we aimed to identify the predominant emotion and intensity associated with specific sounds. Our primary focus was on sounds that evoke medium-to-high levels of disgust; we included three categories of such sounds to tease apart what other emotions they may evoke and the intensity of these emotions. We hypothesized that individuals with misophonia compared to controls would report greater emotion intensities to these sounds, specifically emotions of disgust, anger and annoyance. Concurrently, we monitored changes in pupil diameter while participants listened to the selected sounds. We also examined whether there are distinguishable pupillary responses between emotional categories of sounds. Our intent was to find pupillary feature(s) (e.g., time-to-maximum pupil diameter) that are unique to the experience of feeling disgust in response to a sound. Given prior indications that the pupillary response to visually disgusting stimuli is constriction or eye narrowing (Ayzenberg, Hickey and Lourenco, 2018; Lee and Anderson, 2017), we asked whether the pupillary response to auditorily disgusting stimuli would be smaller than for other emotional sounds. We hypothesized that disgusting sounds would evoke larger changes in pupil diameter for individuals with misophonia compared to controls. For other sound categories, we predicted no such differences. Lastly, we aimed to understand the relationship between reported intensity of an emotion and changes in pupil diameter for both populations. We predicted that emotion intensity would scale linearly with changes in pupil diameter. Two potential outcomes were considered. First, it is possible that misophonic individuals may report greater emotion intensity and have greater pupillary reactivity to the sounds compared to controls, while falling on the same linear function. Alternatively, misophonic listeners may have greater pupillary reactivity for a given level of self-reported emotion intensity, resulting in a linear function positioned above that of controls. Together, these findings could provide support for pupillometry being a useful objective measure to systematically and accurately classify individuals as misophonic or non-misophonic listeners. Furthermore, our paradigm presented the sounds to listeners twice. We expected that the changes in pupil diameter, and reported emotion intensity would increase for both misophonic and non-misophonic listeners due to the phenomenon of mere exposure (Zajonc, 1968), but increases would be reliably larger for misophonic listeners. For both negative and positive sounds, we predicted that the magnitude of pupillary response would increase upon second exposure, while the reported emotion intensity would increase in the direction of the sound's valence. Thus, negative sounds would be rated as more negative, and positive sounds would be rated as more positive upon second exposure; a differential mere exposure effect observed between positive and negative stimuli (see Brickman et al., 1972 and Heller et al., 2024 (*under review*)).

II. Method

Participants

Fifty-nine individuals were recruited ($M_{AGE} = 23.54$; range = 18 to 45 years old, 36 females, 20 males, and one non-binary) from Carnegie Mellon University and the surrounding area in Pittsburgh, PA. Due to known resting and reflexive pupillary differences between young and older adults, we set an age criterion between 18 to 45 years old (Lobato-Rincon et al., 2014; Bitsios, Prettyman, and Szabadi, 1996; Winn et al., 1994; Birren, Casperson, and Botwinick, 1950). Our final sample size was 57 participants. All participants completed the Duke-Vanderbilt Misophonia Screening Questionnaire (DVMSQ) (Williams, Cassiello-Robbins, and Anand, 2022), the Misophonia Questionnaire (MQ) (Wu et al., 2014), and the S-Five (Vitoratou et al., 2023) to assess the severity of sound intolerance. We followed the scoring guidelines provided for each of the scales to categorize participants into 'misophonic' and 'non-misophonic' participant groups. Participants ($N = 14$) in the 'misophonic' category reported symptoms that corresponded to clinical presentations of misophonia.

Participants ($N = 43$) in the 'non-misophonic' category reported symptoms that corresponded to sub-clinical and non-clinical presentations of misophonia.

Although all our participants reported normal hearing, for individuals in the misophonic category, we measured hearing sensitivity and collected loudness discomfort level (LDL) judgements using pure-tone audiometry (MAICO MA41 system). Previous research suggests that clinical presentations of misophonia are often comorbid with decreased sound level tolerance, i.e., hyperacusis (Jastreboff and Jastreboff, 2002). Audiograms confirmed that a majority of our participants had normal hearing up to 8 kHz, defined by the American Speech-Language Hearing Association (ASHA) as hearing thresholds under 20 dB HL (American Speech-Language Hearing Association, 2024). There were four participants who had normal hearing up to 4 kHz with mild hearing loss from 6 to 8kHz. Furthermore, LDL judgements indicated that 11 misophonic participants had some form (mild-to-severe) of sound level intolerance (Sherlock and Formby, 2005; Goldstein and Shulman, 1996) with an average discomfort threshold of 76 dB HL.

The experimental protocol was approved by the CMU Institutional Review Board. Participants signed in-person consent forms and were compensated with course credit or payment.

Sounds and Measurement Scales

Sounds were compiled across three databases: 1) the International Affective Digitized Sounds (IADS) database ($N = 45$) (Bradley and Lang, 2007), 2) the Environmental Sound Classification database (ESC-50) ($N = 3$) (Piczak, 2015) and 3) recordings from our laboratory ($N = 25$). We conducted online preliminary testing to select a final set of sounds which primarily evoked one of the following emotions: happiness, sadness, anger, fear, and disgust. In the pilot study, participants were instructed to listen to and rate a single sound on nine unique scales. Five of the nine scales assessed a single emotion (i.e., *happiness*, *sadness*, *anger*, *fear*, and*disgust*) on a 9-point Likert scale (i.e., 'How ______ did the sound make you feel?') from 1 to 9, where 1 represented not at all feeling a particular emotion (e.g., 'Not at all *happy*'), and 9 represented extremely feeling a particular emotion (e.g., 'Extremely*happy*'). The remaining four scales assessed: 1) *arousal*, using a 9-point bi-polar scale (i.e., 'How calm or excited did this sound make you feel?') from very calm (1) to very excited (9), 2)*valence*, using a 9-point bi-polar scale (i.e., 'How negative or positive was this sound?' from very negative (1) to very positive (9), 3) *annoyance*, using an 11-point Likert scale (i.e., 'How much are you bothered, disturbed or, annoyed by this sound?') based on the ISO standard (i.e., ISO_TS_-15666_2021), ranging from "not at all feeling bothered, disturbed, or annoyed" (0) to "extremely feeling. . . ." (10), and lastly, 4) *pleasantness*, using an 11-point bi-polar scale (i.e., 'How pleasant is the sound?') (Heller and Smith, 2022), with -5 representing a very unpleasant sound, 0 representing a neutral sound, and +5 representing a very pleasant sound. Based on preliminary testing, we selected a final set of 26 sounds. There were between three and five sounds that represented each emotion category (refer to Table 1). We curated a 'misophonic trigger' sound category based on other preliminary lab data that showed differences in pleasantness ratings between misophonic and non-misophonic listeners for a set of sounds. Note, these sounds have been anecdotally reported as being 'triggers' for some individuals with clinical levels of misophonia. All sound levels were normalized based on root-mean square (RMS) value to ensure that each sound was presented within a similar range of perceptual loudness (Auditory Toolbox; Mathworks, 2022a). This type of normalizing procedure has been reported to mitigate loudness confounds for affective responses, such as pleasantness (Oszczapinska et al., 2024). All of the sounds had a continuous stream of events between one and six seconds ($M_{\text{TOTAL DURATION}} = 5.58$ seconds, Range = 3.35 to 6 seconds).

Table 1. Describes the final set of sounds by detailing the*a priori* emotion category they were assigned to based on preliminary data, and the database they were borrowed from. The duration column indicates the total duration of the sound in seconds. If the sound file contained multiple events, Duration is the total time accounting for onsets and offsets of events.

Sound	***A priori* Emotion Category**	**Database**	**Duration (seconds)**	**# o**
Young children playing outside	Happy-Excited	IADS	6	1
Applause from a large audience	Happy-Excited	IADS	6	1

Positively cheering sports crowd	Happy-Excited	IADS	6	1
Mountain stream bubbling	Happy-Calm	IADS	6	1
Mountain stream flowing	Happy-Calm	IADS	3.58	1
Valley ambiance (cave dripping with birds)	Happy-Calm	IADS	6	1
Distressed infant crying	Sad	IADS	5.68	1
Man sobbing	Sad	IADS	6	1
Woman crying	Sad	IADS	6	1
Crow cawing	Fear	IADS	4.77	4
Swarm of bees/insects buzzing	Fear	IADS	6	1
Large animal (bear) growling	Fear	IADS	6	1
Car-crashing scene	Fear	IADS	4.27	1
Fake belching	Disgust-Medium	IADS	1.46	2
Male sneezing	Disgust-Medium	IADS	5.71	1
Female coughing	Disgust-Medium	IADS	6	1
Male sipping a beverage through a straw	Disgust-Medium	IADS	6	1
Real belching	Disgust-High	IADS	3.35	2
Person chewing gum (wet chewing)	Disgust-High	IADS	6	1
Female coughing	Disgust-High	IADS	6	1
Individual vomiting	Disgust-High	IADS	6	1
Car engine starting	Misophonic trigger	ESC-50	6	1
Door knocking	Misophonic trigger	ESC-50	2.81	1
Female eating chips (crunchy-chewing)	Misophonic trigger	Lab recording	5.1	1
Utensils scraping against one another	Misophonic trigger	Lab recording	6	1
Water dripping from a faucet	Misophonic trigger	ESC-50	1.2	2

jabbrv-ltwa-all.ldf jabbrv-ltwa-en.ldf

Procedure

The experiment took place in-person within a sound attenuating chamber covered with sound absorbing foam ($M_{\text{AMBIENT LUMINANCE}}$ = 66 Lux). Sound presentation was controlled with a custom MATLAB program (MathWorks, 2022a) on a 14-inch Dell monitor (Latitude 5421, 1920 x 1080 pixels, 60 Hz refresh rate). Participants sat approximately 60 cm away from the laptop monitor. The sounds were presented twice, in random order, over the course of one 40-minute session through Koss UR20 over-ear headphones at 62 dB SPL. Pupillary responses were measured using the SmartEye Aurora eye-tracker (Version SE XO v9.3) recording at 30 Hz. The SmartEye Aurora has tracking capabilities that allow the pupil to be accurately measured without the need of a chin rest.

Each trial began with two seconds of a blank, gray screen with RGB values of: 108, 108, 108; $M_{\text{LUMINANCE}}$ = 40 cd/m^2). Next, a black fixation cross (RGB: 0, 0, 0) was presented in the center of the gray screen. The fixation cross appeared two seconds prior to the presentation of the sound, and remained until the sound ended. Listeners were instructed to maintain their gaze on the fixation cross for its entire duration. Each sound clip had the length of six seconds regardless of whether the sound event filled the entire time. There was a one-second delay between the offset of the sound and the beginning of the rating task. Listeners completed the same set of judgments as described in the pilot study (see section 'B. Sounds and Measurement Scales). For *valence* and*arousal* scales, we instructed listeners to use '5' as the midpoint. For the *pleasantness* scale, we instructed listeners to use '0' as the midpoint. After data collection had already begun, we included one additional question which asked the listener how much they wanted to escape from the sound, using a 10-point Likert scale, with 0 representing not at all wanting to escape, and 10 representing extremely wanting to escape the sound. After rating each sound on the 10 unique scales, listeners were prompted to press the 'SPACE' button to proceed to the next sound.

Pupillary pre-processing

Pupil diameter measurements were pre-processed using a custom MATLAB program (MathWorks, 2023a) for each individual sound at the participant level. For more details, please refer to Supplementary Material and Supplemental Figure 1. To maintain high data quality, we removed individual participant trials that had more than 50% of missing data as well as those who had missing data within a 500-ms time window prior to the start of the sound. Note, we used the mean pupillary value across the 500 ms time window prior to sound presentation as our baseline for future subtraction (Winn et al., 2018). This resulted in the removal of 448 individual trials (448/2964, 15.1% of total data); even with this removal, the average number of contributed samples and number of participants remained equal and consistent across sounds (Supplemental Table 1). When visualizing changes in pupil diameter in our figures, we utilized a linear interpolation method to estimate the change in pupil diameter across regions that no longer had data due to blink or artifact removal (Hershman, Henik and Cohen, 2018). Otherwise, analysis on pupillary changes used data without any linear interpolation.

To examine pupillary changes across time for each sound, we subtracted the pupil diameter during sound presentation from the pupil diameter prior to sound presentation (baseline) at the participant level. Then, we sub-averaged trials for individual sounds across participants. For those analyses which investigated pupillary differences across sound categories, we calculated a sub-average across sounds in a single emotion category (refer to Table 1).

III. Results

Behavioral outcomes

There were three stages of analyzing self-reported emotional intensity ratings. First, we validated that the selected sounds evoked the highest intensity congruent with the *a priori* emotion category they had been assigned to based on preliminary testing. Second, we examined whether differences existed in emotional intensity ratings between misophonic and non-misophonic listeners. Regarding the specific emotions of disgust, anger, and annoyance felt in response to both *a priori* misophonic trigger sounds and disgusting sounds (i.e., Disgust-High and Disgust-Medium), we predicted more intense emotions for misophonic than non-misophonic listeners. Additionally, we predicted that misophonic listeners would report greater arousal and feelings of wanting to escape sounds toward misophonic trigger sounds and disgusting sounds compared to non-misophonic listeners. Finally, we investigated whether a second exposure to each sound would produce consistent emotional intensity ratings for either or both populations. Prior to analysis, we converted and equated all of the scales to a 9-point Likert scale with values ranging from 1 to 9. In this case, we re-coded judgments of sound pleasantness, annoyance, and wanting to escape to continuous, equidistant values from 1 to 9. For *arousal* and*valence* scales, we converted the two bipolar scales into four unipolar scales: negative, positive, calmness and excitement, all ranging from 1 to 9. To do this, we took all raw responses from 1 to 5 as belonging to one of the two new scales (i.e., 'negative' and 'calmness') with every response greater than 5 re-coded as a 5. A '5' in this case represented a 'neutral' point of not feeling an emotion. This was done because pupil diameter is shown to increase for both positive and negative valence sounds. To be congruent with all other scales going from low to high intensity of an emotion, these responses were transformed by taking the absolute value of the subtraction (i.e., abs(1-6) = 5). Raw responses from 5 to 9 belonged to the other companion scale (i.e., 'positive' and 'excitement') with every response less than 5 re-coded as a 5. Each of the new scales were re-coded to values ranging from 1 to 9. Responses from all 57 participants were included.

Validating *a priori* emotion categories

To confirm that the sounds in each category evoked the highest intensity congruent with their *a priori* emotion category (refer to Table 1), we examined whether the average emotion rating for a specific emotion category was reliably higher for sounds in that category compared to sounds in other categories. We used the 'emmeans 'package in R (Lenth, 2024) to conduct contrast analyses with a Tukey adjustment, separately for first and second sound exposure.

Sounds evoking Fear. The sounds assigned to the 'Fear' emotion category (M_{FEAR} = 4.17, 95% CI = [3.81, 4.53]) reliably evoked a greater intensity of feeling *afraid*compared to all other sounds (Mean range$_{OTHER}$ = 1.49 to 2.87). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure (M_{FEAR} = 4.02, 95% CI = [3.66, 4.38]; Mean range$_{OTHER}$ = 1.42 to 2.60).

Sounds evoking Disgust. Disgusting sounds were assigned to one of two *a priori* 'Disgust' emotion categories based on the level of disgust they evoked in the pilot study. These categories were separated in advance in anticipation of the need to make comparisons across emotion categories that were matched in valence. The sounds in the 'Disgust-High' emotion category ($M_{DISGUST\text{-}HIGH}$ = 6.61, 95% CI = [6.22, 7.01]) reliably evoked a greater intensity of feeling *disgust* compared to all other sounds, including sounds in the 'Disgust-Medium' emotion category (Mean range$_{OTHER}$ = 1.19 to 4.88). The comparisons consistently showed significant differences at or below a significance level of $p = 0.001$, even following Tukey correction. Similarly, the sounds in the 'Disgust-Medium' emotion category ($M_{DISGUST\text{-}MEDIUM}$ = 4.88, 95% CI = [4.49, 5.28]) reliably evoked a greater intensity of feeling *disgust* compared to all other sounds, excluding sounds in the 'Disgust-High' category. The comparisons consistently showed significant differences at or below a significance level of p = 0.001, even following Tukey correction. Sound ratings for both categories were consistent upon second exposure ($M_{DISGUST\text{-}HIGH}$ = 6.38, 95% CI = [5.99, 6.78]; Mean range$_{OTHER}$ = 1.22 to 4.73) ($M_{DISGUST\text{-}MEDIUM}$ = 4.73, 95% CI = [4.33, 5.12]).

Sounds evoking Happiness. Happy sounds were assigned to one of two *a priori* 'Happy' emotion categories based on the level of arousal they evoked in the pilot study. The sounds in the 'Happy-Excited' emotion category (i.e., happy excitement) ($M_{HAPPY\text{-}EXCITED}$ = 5.39, 95% CI = [5.07, 5.70]) reliably evoked a greater intensity of feeling *happy* compared to all other sounds, including sounds in the 'Happy-Calm' category (i.e., happy calmness) (Mean range$_{OTHER}$ = 1.37 to 4.51). The comparisons consistently showed significant differences below a significance level of $p = 0.001$, even following Tukey correction. The sounds in the 'Happy-Calm' emotion category ($M_{HAPPY\text{-}CALM}$ = 4.51, 95% CI = [4.20, 4.82]) reliably evoked a greater intensity of feeling *happy* compared to all other sounds, except for sounds in the 'Happy-Excited' category. The comparisons consistently showed significant differences at or below a significance level of $p = 0.001$, even following Tukey correction. Sound ratings for both categories remained consistent upon second exposure ($M_{HAPPY\text{-}EXCITED}$ = 5.02, 95% CI = [4.71, 5.33]; Mean range$_{OTHER}$ = 1.30 to 4.27) ($M_{HAPPY\text{-}CALM}$ = 4.27, 95% CI = 3.96, 4.59]).

Sounds evoking Sadness. The sounds assigned to the 'Sad' emotion category (M_{SAD} = 4.28, 95% CI = [3.94, 4.62]) reliably evoked a greater intensity of feeling *sad*compared to all other sounds (Mean range$_{OTHER}$ = 1.40 to 2.65). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure (M_{SAD} = 3.97, 95% CI = [3.63, 4.31]; Mean range$_{OTHER}$ = 1.37 to 2.23).

Supplementary ratings of a priori sound categories

Arousal. The sounds in the 'Happy-Excited' category were found to evoke the greatest amount of *excitement* receiving an average rating of 3.61 (95% CI = [3.24, 3.98]), compared to all other sounds (Mean range$_{OTHER}$ = 1.63 to 2.75). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{HAPPY\text{-}EXCITED}$ = 3.83, 95% CI = [3.46, 4.20]; Mean range$_{OTHER}$ = 1.47 to 2.78). In contrast, the sounds in the 'Happy-Calm' category were found to evoke the greatest amount of*calmness*, receiving an average rating of 5.00 (95% CI = [4.57, 5.43]), compared to all other sounds (Mean range$_{OTHER}$= 1.76 to 2.38). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{HAPPY\text{-}CALM}$ = 5.05, 95% CI = [4.61, 5.48]; Mean range$_{OTHER}$ = 1.72 to 2.38).

Valence. The sounds in the 'Disgust-High' category were perceived as the most *negative*, receiving an average rating of 6.70 (95% CI = [6.36, 7.04]) compared to all other sounds, except for sounds in the 'Sad' category

(M_{SAD} = 6.13, 95% CI = [5.76, 6.51]; Mean $\text{range}_{\text{OTHER}}$ = 1.29 to 4.80). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{DISGUST-HIGH}}$ = 6.42, 95% CI = [6.08, 6.76]; M_{SAD} = 5.92, 95% CI = [5.55, 6.30]; Mean $\text{range}_{\text{OTHER}}$ = 1.31 to 3.50). In contrast, the sounds in the 'Happy-Excited' category were perceived as the most *positive*, receiving an average rating of 5.32 (95% CI = [5.07, 5.56]), compared to all other sounds (Mean $\text{range}_{\text{OTHER}}$ = 1.07 to 4.15). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{HAPPY-EXCITED}}$ = 4.98, 95% CI = [4.73, 5.22]; Mean $\text{range}_{\text{OTHER}}$ = 1.04 to 4.15).

Pleasantness. The sounds in the 'Disgust-High' category were perceived as the most *unpleasant*, receiving an average rating of 2.01 (95% CI = [1.80, 2.22]), compared to all other sounds (Mean $\text{range}_{\text{OTHER}}$ = 2.52 to 6.70). The comparisons consistently showed significant differences at or below a significance level of $p <$ 0.001, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{DISGUST-HIGH}}$ = 2.14, 95% CI = [1.93, 2.35]; Mean $\text{range}_{\text{OTHER}}$ = 2.80 to 6.70). The sounds in the 'Happy-Calm' category were perceived as the most *pleasant*, receiving an average rating of 6.70 (95% CI = [6.46, 6.94]), compared to all other sounds except 'Happy-Excited' sounds ($M_{\text{HAPPY-EXCITED}}$ = 6.57, 95% CI = [6.44, 6.91],p = 1.0; Mean $\text{range}_{\text{OTHER}}$ = 2.01 to 3.90). The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{HAPPY-CALM}}$ = 6.70, 95% CI = [6.46, 6.94]; $M_{\text{HAPPY-EXCITED}}$ = 6.40, 95% CI = [6.16, 6.64],p = 1.00; Mean $\text{range}_{\text{OTHER}}$ = 2.14 to 3.85).

We observed high correlations for both the *negative* valence and*unpleasant* scales (r = -0.96, $t(24)$ = -16.29,p $<$ 0.001), and for the *positive* valence and*pleasant* scales (r = 0.92, $t(24)$ = 11.91, $p<$ 0.001). The obvious neutral point of zero on the*pleasantness* scale aligned with the less obvious neutral point of 5 on the *valence* scale, validating our choice to split up the valence scale at 5. Since sound valence is equivalently captured by both the *pleasantness* and *valence* scales, we only included participant ratings from *negative* and *positive* valence scales in future analyses.

Annoyance. The sounds in the 'Disgust-High' emotion category ($M_{\text{DISGUST-HIGH}}$ = 6.56, 95% CI = [6.16, 6.97]) reliably evoked the greatest amount of *annoyance*compared to all other sounds. In contrast, sounds in the 'Happy-Excited' category ($M_{\text{HAPPY-EXCITED}}$ = 1.82, 95% CI = [1.40, 2.24]) and 'Happy-Calm' category ($M_{\text{HAPPY-CALM}}$ = 2.01, 95% CI = [1.58, 2.43], p = 0.96) reliably evoked the least amount of *annoyance* compared to all other sounds. The comparisons consistently showed significant differences below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{DISGUST-HIGH}}$ = 6.28, 95% CI = [5.88, 6.68]; $M_{\text{HAPPY-EXCITED}}$ = 2.01, 95% CI = [1.59, 2.43]; $M_{\text{HAPPY-CALM}}$ = 1.87, 95% CI = [1.45, 2.30], p= 0.99).

Anger. The sounds in the 'Disgust-High' category ($M_{\text{DISGUST-HIGH}}$ = 4.13, 95% CI = [3.76, 4.49]) reliably evoked the greatest amount of *anger* compared to all other sounds, including sounds in the 'Disgust-Medium' category (Mean $\text{range}_{\text{OTHER}}$ = 1.40 to 3.28). In contrast, sounds in the 'Happy-Excited' category ($M_{\text{HAPPY-EXCITED}}$ = 1.40, 95% CI = [1.02, 1.79]) and 'Happy-Calm' category ($M_{\text{HAPPY-CALM}}$ = 1.43, CI = [1.05, 1.82], p = 1.0) reliably evoked the least amount of *anger.* The comparisons consistently showed significant differences at or below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{DISGUST-HIGH}}$ = 4.03, 95% CI = [3.66, 4.39]; $M_{\text{HAPPY-EXCITED}}$ = 1.35, 95% CI = [0.97, 1.74]; $M_{\text{HAPPY-CALM}}$ = 1.41, 95% CI = [1.02, 1.80], p= 1.0).

Escape. Listeners expressed the greatest desire to*escape* the sounds in the 'Disgust-High' category ($M_{\text{DISGUST-HIGH}}$ = 6.64, 95% CI = [6.09, 7.18]). In contrast, listeners were least inclined to *escape* sounds in the 'Happy-Calm' category ($M_{\text{HAPPY-CALM}}$ = 2.07, 95% CI = [1.50, 2.65]) and 'Happy-Excited' category ($M_{\text{HAPPY-EXCITED}}$ = 2.14, 95% CI = [1.57, 2.71],p = 1.0) compared to all other sounds. The comparisons consistently showed significant differences below a significance level of $p < 0.001$, even following Tukey correction. Sound ratings were consistent upon second exposure ($M_{\text{DISGUST-HIGH}}$ = 6.39, 95% CI = [5.85, 6.93]; $M_{\text{HAPPY-CALM}}$ = 1.94, 95% CI = [1.37, 2.52]; $M_{\text{HAPPY-EXCITED}}$ = 2.18, 95% CI = [1.61, 2.75],p = 0.97).

Note: Unlike all the other self-report scales, we only collected responses from 37 ($N_{\text{MISO}} = 13, N_{\text{NON-MISO}} = 24$) out of 57 participants because the question was added to the experimental design after data collection had already begun.

Differences in emotional intensity ratings between misophonic and non-misophonic listeners

One of our primary objectives was to investigate whether there were measurable differences in emotional intensity ratings between misophonic and non-misophonic listeners. Using the 'lme4' package (Bates et al., 2015) in R (R Core Team, 2024; version 4.3.3), we constructed a linear mixed-effects model with exposure (first or second), *a priori*emotion category of the sounds (e.g., happy-excited), emotion scale (e.g., *disgust*, *valence*) and misophonic status (misophonic or non-misophonic) as fixed effects, and participant number as a random effect to account for intercept differences. The model estimated the difference in emotion intensity ratings reported by non-misophonic and misophonic listeners. Type-3 ANOVA statistics were obtained using the 'lmerTest' package. Both F and T test statistics and their degrees of freedom were computed using Satterthwaite's method (Kuznetsova et al., 2017). Note, we excluded the 'escape' emotion intensity scale and corresponding ratings from our model given that we did not collect responses from all our listeners. At the end of this section, we examined 'escape' ratings in a separate analysis.

We observed a main effect of exposure ($F(1, 29305) = 14.97$, $p < 0.001$), indicating that all emotion intensity ratings were significantly higher upon first exposure to the sounds. Significant main effects of emotion intensity scale ($F(9, 29305) = 332.98$, $p < 0.001$), and *a priori*emotion category of the sounds ($F(6, 29305) = 136.18, p < 0.001$) showed that reported emotion intensity was greater on some scales and the intensity of a particular emotion differed between *a priori* sound categories while a two-way interaction between emotion intensity scale **X** *a priori*emotion category ($F(54, 29305) = 158.91$, $p< 0.001$), showed that the scales receiving the highest rating depended on the category of the sound. Further, we observed interactions with misophonic status: 1) emotion intensity scale **X** misophonic status ($F(9, 29305) = 29.27$, $p < 0.001$), and 2) *a priori* emotion category **X** misophonic status ($F(6, 29305) = 4.69$, $p < 0.001$) and 3) a three-way interaction between emotion intensity scale **X**misophonic status **X** *a priori* emotion category ($F(54, 29305) = 3.14$, $p < 0.001$). These interactions show that differences in emotion intensity ratings between our two populations vary across the scales and sound categories. We found no interaction with exposure, suggesting that differences in emotion intensity ratings between non-misophonic and misophonic listeners were consistent across first and second sound exposure.

The three-way interaction between emotion intensity scale, misophonic status and *a priori* emotion category was followed up with focused post-hoc t-tests on estimated marginal means ('emmeans' in R; Lenth, 2024) to reveal differences in emotion intensity ratings between non-misophonic and misophonic listeners within *a priori* sound categories. Figure 1 illustrates the average emotion intensity ratings reported by non-misophonics (solid bars) and misophonics (stripped bars) for misophonic trigger sounds, and for sounds that evoke medium and high levels of disgust.

For sounds in the *a priori* 'Disgust-High' category, misophonic listeners reported significantly greater intensity of *disgust*($M_{\text{MISO}} = 7.17$, $M_{\text{NON}} = 6.28$, *t-ratio*$(232) = -3.12$, $p < 0.01$), *anger*($M_{\text{MISO}} = 4.98$, $M_{\text{NON}} = 3.78$, *t-ratio*$(232) = -4.20$, $p < 0.001$), *annoyance* ($M_{\text{MISO}} = 7.03$, $M_{\text{NON}} = 6.22$, *t-ratio*$(232) = -2.82$, $p < 0.01$), and *fear* ($M_{\text{MISO}} = 3.74$, $M_{\text{NON}} = 2.27$, *t-ratio*$(232) = -5.14$, $p < 0.001$) compared to non-misophonic listeners. We also observed that misophonic listeners perceived the sounds in this category to have greater negative valence than non-misophonic listeners ($M_{\text{MISO}} = 7.12$, $M_{\text{NON}} = 6.38$, *t-ratio*$(232) = 2.62$, $p< 0.01$). These rating trends were similar for sounds in the *a priori* 'Disgust-Medium' category. Misophonic listeners report significantly greater intensity of *disgust* ($M_{\text{MISO}} = 5.38$, $M_{\text{NON}} = 4.62$, *t-ratio*$(232) = -2.65, p < 0.01$), *anger* ($M_{\text{MISO}} = 3.75$, $M_{\text{NON}} = 2.96$, *t-ratio*$(232) = -2.76$, $p< 0.01$), and *fear* ($M_{\text{MISO}} = 3.04$, $M_{\text{NON}} = 2.07$, *t-ratio*$(232) = -3.38$, $p< 0.001$) compared to non-misophonic listeners. We also observed a reliable difference in reports of how *negative* the sounds in the 'Disgust-Medium' category were perceived by the two populations of listeners ($M_{\text{MISO}} = 5.27$, $M_{\text{NON}} = 4.66$, *t-ratio*$(186) = -2.12$, $p = 0.04$).

Figure 1. The average emotion intensity rating reported by misophonic listeners (striped bars) compared

to non-misophonic listeners (solid bars) for sounds in *a priori* categories of disgust (i.e., 'Disgust-High' and 'Disgust-Medium'), and misophonic trigger sounds ('Triggers'). Each panel of data, delineated by the header at the top of each individual graph, corresponds to a single emotion scale: *disgust, anger*, *annoyance*, *fear*,*calmness (arousal), excitement (arousal), negative (valence)* and*positive* (valence). The error bars correspond to the 95% confidence interval of the estimated marginal means.

For sounds in the *a priori* misophonic triggers category ('Triggers'), we observed that misophonic listeners report reliably greater intensity of *disgust* (M_{MISO} = 3.64, M_{NON} = 2.64, *t-ratio*(186) = -3.69, $p<$ 0.001), *anger* (M_{MISO} = 3.83, M_{NON} = 2.83, *t-ratio*(186) = -2.86, $p<$ 0.01), *annoyance* (M_{MISO} = 4.85, M_{NON} = 3.89, *t-ratio*(186) = -3.54, $p<$ 0.001), and *fear* (M_{MISO} = 2.96, M_{NON} = 2.17, *t-ratio*(186) = -2.96, $p<$ 0.01) compared to non-misophonic listeners. Furthermore, we observed that misophonic listeners report feeling significantly less*calm* compared to misophonic listeners (M_{MISO} = 1.73, M_{NON} = 2.59, *t-ratio*(186) = 3.18, $p<$ 0.01) when listening to these sounds, but also greater*excitement* (arousal) compared to non-misophonic listeners (M_{MISO} = 2.61, M_{NON} = 2.04,*t-ratio*(186) = 2.12, p = 0.04). Similar to ratings for sounds in the disgust category, misophonic listeners reliably perceive the sounds in the Triggers category to have greater *negative*valence (M_{MISO} = 4.09, M_{NON} = 3.15,*t-ratio*(186) = 3.46, p < 0.001).

Given that our primary focus is to investigate differences for specific categories of sounds, we refrain from reporting differences in emotion intensity ratings for all other categories of sounds. These additional comparisons can be easily accessed and found in supplementary material in Table 2. In short, we observe that misophonic listeners reported greater intensity of emotions for most sounds compared to non-misophonic listeners.

Reports of wanting to 'escape' the sound. As noted above, 37 of our 57 listeners had the opportunity to provide a response to our 'escape' question that inquired about how much the listener wanted to escape from the sound upon presentation. We conducted a separate analysis on the responses we collected to the *escape* scale for each of our sound categories of interest: sounds in the *a priori*'Disgust-High' category, sounds in the *a priori* 'Disgust-Medium' category, and sounds in the *a priori* 'Triggers' category.

For sounds in the *a priori* 'Disgust-High' category, misophonic listeners (M_{MISO} = 6.81) did not report a greater desire to escape the sounds (M_{NON} = 6.35, t(52) = -0.87,p = 0.39). We observed similar trends for sounds in the *a priori* 'Disgust-Medium' category (M_{MISO} = 5.26, M_{NON} = 4.83, t(52) = -0.84, p = 0.41), and for sounds in the 'Triggers' category (M_{MISO} = 4.72, M_{NON} = 4.04, t(47) = -1.33, p = 0.19). All other comparisons can be found in Table 3 in supplementary material. Given that we did not observe reliable differences in desire of wanting to escape these classes of sounds between our two populations, we believe that the *escape* scale lacks sensitivity for identifying misophonia severity. On the basis of this result, we did not move forward with analyzing the responses to the *escape* scale in the remainder of the results.

Pupillary outcomes

Our time series analysis consisted of four measures that concisely capture the types of pupillary changes that may occur across the duration of a sound. For each of our measures, we predicted that misophonic listeners would have a significantly greater change in pupil diameter in response to the sounds that evoke disgust compared to non-misophonic listeners. First, we examined the mean response at all time points and its associated confidence interval. We chose to focus on the time interval from 2-to-3 seconds because that region showed the maximum effect size between misophonic and non-misophonic listeners. Second, we examined the difference in magnitude of peak pupil response between misophonic and non-misophonic listeners. Third, we analyzed the difference in time course to peak pupil response between misophonic and non-misophonic listeners. Lastly, we investigated whether the change in pupil diameter sustains upon sound offset.

Despite our efforts to ensure comparable valence intensity and arousal levels across sound categories, we were unable to achieve complete equalization. This is a challenging goal given the inherent diversity of our sounds. Consequently, we acknowledge the limitations in conducting direct comparisons of pupillary signatures between sound categories. Furthermore, the absence of significant pupillary differences between our populations for any category may be attributed to underlying variations in the general dimensions of arousal and valence rather than being a difference between specific emotions.

Average change in pupil diameter across time

Using the R Statistical language package 'lme4', we constructed a linear mixed-effects model with exposure (first or second), *a priori*emotion category of the sounds (e.g., happy-excited) and misophonic status (misophonic or non-misophonic) as fixed effects, and participant number as a random effect to account for intercept differences. We used these variables to estimate the difference in the average change in pupil diameter from baseline across six seconds in one-second time intervals between misophonic and non-misophonic listeners. The analysis reported below corresponds to the results for the time interval from 2-to-3 seconds.

From our model including both exposures, we observed main effects of exposure, indicated by a greater change in pupil diameter at first versus second sound exposure ($F(1, 2428) = 68.42$, $p < 0.001$), and *a priori* emotion category of the sounds ($F(6, 2425) = 12.54$, $p < 0.001$), demonstrated by a greater change in pupil diameter for sounds in the 'Disgust-Medium' and 'Triggers' categories compared to sounds in the 'Happy-Calm' category, for example. We did not observe a main effect of misophonic status ($F(1, 55) = 1.68$, $p = 0.20$). We observed a single, significant misophonic status **X** exposure two-way interaction ($F(1, 2428) = 5.59$, $p = 0.02$) shown by a greater pupil response from misophonic listeners upon first exposure to sounds. Due to this interaction, we followed up with separate analyses for first and second exposures. We found no other interactions.

Figure 2 illustrates the results from first exposure. The average change in pupil diameter across time (in seconds) collapsed across all the sounds for misophonic (dashed line) and non-misophonic (solid line) listeners. A positive change indicates that the pupil diameter has increased (dilated) with respect to baseline. Overall, misophonic listeners have greater pupillary dilation in response to the emotional sounds compared to non-misophonic listeners, evidenced by the dashed line remaining above the solid line for most of the sound duration. This difference in pupil dilation was reliable at the 2-to-3 second time interval ($M_{MISO} = 0.30$ mm, $M_{NON} = 0.21$ mm; *t-ratio*(133) = -2.08, $p = 0.04$) and extended through the 5-to-6 second time interval ($M_{MISO} = 0.22$ mm, $M_{NON} = 0.13$ mm; *t-ratio*(133) = -2.01, $p = 0.047$). Upon second exposure to the sounds (not shown), the mean pupil diameter reliably decreased for misophonic (change = -0.12 mm, $p < 0.001$), and non-misophonic listeners (change = -0.06 mm, $p < 0.001$), respectively. However, the difference in average pupil diameter no longer reached significance ($M_{MISO} = 0.18$ mm, $M_{NON} = 0.15$ mm, *t-ratio*(127) = -0.60, $p = 0.55$). This was true for all other time intervals as well (see Supplemental Table 4).

Figure 2. The average change in pupil diameter (mm) relative to baseline across time (seconds) during first exposure. The plot contains separate functions for misophonic (dashed line) and non-misophonic (solid line) listeners collapsed across all sounds. The shaded regions denote 95% confidence intervals. An upward deflection indicates that the pupil has dilated compared to baseline.

In Figure 3, we illustrate the average change in pupil diameter relative to baseline across time, upon first exposure, for each *a priori*sound category. At 2-to-3 seconds, misophonic listeners had reliably greater pupil dilation relative to baseline (M = 0.36 mm) from non-misophonic listeners (M = 0.22 mm) *only* for sounds in the 'Triggers' category (*t-ratio*(154) = -2.65, $p = 0.01$). However, we did observe that there was a difference in pupil diameter for 'Disgust-Medium' sounds that reached marginal significance ($M_{MISO} = 0.35$ mm, $M_{NON\text{-}MISO} = 0.23$ mm; *t-ratio*(187) = -1.99, $p = 0.05$) at this same time point. The difference in pupil dilation for Triggers remained reliable up to the 4-to-5 second time interval. At this point, we observed reliable differences in pupil dilation between our populations for sounds in the 'Disgust-High' category ($M_{MISO} = 0.36$ mm, $M_{NON} = 0.21$ mm; *t-ratio*(209) = -2.19, $p = 0.03$) which sustained until the end of the sound. These results can be found in Supplemental Table 5. Upon second exposure, provided in Supplemental Table 6, we no longer observed a reliable difference in pupil diameter between our populations for any category of sounds at any time interval.

Figure 3. The average change in pupil diameter (mm) relative to baseline across time (seconds) during first exposure. The plot contains separate functions for misophonic (dashed line) and non-misophonic (solid

line) listeners. Each panel contains two functions for a single *a priori* sound category (e.g., 'Disgust-High', 'Happy-Excited'). The shaded regions denote 95% confidence intervals. An upward deflection from zero indicates that the pupil has dilated compared to baseline.

Differences in magnitude of maximum change in pupil diameter

Using the R (R Core Team, 2024; version 4.3.3) package 'lme4' (Bates et al., 2015), we constructed a linear mixed-effect model with the same variables described in *section B.A* above. In this section, we used the variables to estimate the difference in the magnitude of maximum change in pupil diameter between misophonic and non-misophonic listeners. Similarly, we observed that the change in pupil diameter significantly differed between *a priori* sound categories ($F(6,\ 2436) = 9.69$, $p < 0.001$). The greatest change in pupil diameter, taken across all sound categories, occurred upon first exposure ($F(1,\ 2439) = 78.98$, $p < 0.001$). The only significant interaction we observed was between misophonic status **X** exposure ($F(6,\ 2439) = 5.04$,$p =$ 0.02). Thus, reliable differences in the magnitude of maximum change in pupil diameter *may* exist between misophonic and non-misophonic listeners, but magnitude of maximum change in pupil diameter appears to reliably depend on the specific set of sounds and listening time.

Generally, we find that misophonic listeners have greater maximum change in pupil dilation response, collapsed across all sounds, compared to non-misophonic listeners, but the magnitude of the response is not significantly larger ($M_{MISO} = 0.44$ mm, $M_{NON} = 0.38$ mm; *t-ratio*(61) = -1.63, $p = 0.11$). Upon second exposure to the sounds, the magnitude of the response reliably decreases for misophonic (change = 0.11 mm, $p <$ 0.001) and non-misophonic listeners (change = 0.07 mm,$p < 0.001$), respectively. However, the difference in magnitude of maximum change in pupil diameter between the populations remains not significant at second exposure ($M_{MISO} = 0.33$ mm, $M_{NON} = 0.31$ mm, *t-ratio*(61) = -0.56, p= 0.58).

When examining individual emotion categories to which the sounds were assigned, the maximum change in pupil diameter for misophonic listeners ($M_{MISO} = 0.48$ mm) was reliably greater than non-misophonic listeners ($M_{NON} = 0.37$ mm) *only*for sounds in the 'Trigger' category (*t-ratio*(133) = -2.09,$p = 0.04$). In Supplemental Table 7, we observed that upon second exposure, the maximum change in pupil diameter for misophonic listeners was not reliably larger than non-misophonic listeners for any of the sounds.

Differences in latency to maximum change in pupil diameter

We constructed a linear mixed-effect model with the same variables described in *section B.A* to predict differences in latency to maximum change in pupil diameter. Unlike previous findings, we only observe that the time to maximum change in pupil diameter significantly differed between *a priori* sound categories ($F(6,\ 2437) = 3.13$, $p = 0.005$) with the longest latency occurring at first exposure to the sounds ($F(1,\ 2441) =$ 14.00, $p < 0.001$). We did not observe any significant interactions in our model.

When collapsing pupil responses across all sounds, we observed that the time course to maximum change in pupil diameter was slightly longer for misophonic listeners, at 3.43 seconds, compared to non-misophonic listeners, at 3.29 seconds, but the difference in time was not statistically meaningful (*t-ratio*(69) = -0.68, $p =$ 0.50). Upon second exposure, the time course to maximum change in pupil diameter *was* reliably shorter for both populations ($change_{MISO} = 340$ ms, $p < 0.01$; $change_{NON} = 170$ ms, $p = 0.01$). However, there was no reliable difference in latency to maximum change in pupil diameter between misophonic and non-misophonic listeners at second exposure ($T_{MISO} = 3.09$ seconds, $T_{NON} = 3.12$ seconds, *t-ratio*(69) = 0.18, $p = 0.86$).

When considering sound emotion categories separately, we *only*observed a reliable difference in the time course to maximum change in pupil diameter for sounds in the 'Disgust-High' category. For this category of sounds, the maximum change in pupil diameter occurred at 3.99 seconds for misophonic listeners, compared to 3.36 seconds for non-misophonic listeners (*t-ratio*(340) = -2.07, $p = 0.04$). In subsequent exposure, the time course to maximum change in pupil diameter was no longer reliably longer or shorter for misophonic listeners for any *a priori* sound category (refer to Supplemental Table 8).

Change in pupil diameter decreases upon sound offset

In this study, we observed that the maximum change in pupil diameter occurs approximately halfway through sound presentation in a time window of one to six seconds. Following the point of maximum dilation, Figure 2 shows that the change in pupil diameter gradually decreases over time. In other words, the pupil diameter appears to constrict back toward baseline. However, previous research has observed that maximum pupillary response occurs approximately one to one-and-half seconds following stimulus offset (Winn et al., 2018). Therefore, we address this previous finding by comparing the change in pupil diameter at five-to-six seconds, a time point where the sound is coming to an end, to six-to-seven seconds, a time point of one second following sound offset. Additionally, given that sound events have different durations (see Table 1), we compared the change in pupil diameter during the final second of each sound to that of one second following sound offset.

Comparing the change in average pupil diameter at five-to-six second time window to six-to-seven second time window

Collapsing across all sounds, Figure 2 illustrates that the average change in pupil diameter reliably decreased for both misophonic (change = 0.02 mm, $t(25) = 2.20$, $p = 0.04$) and non-misophonic (change = 0.04 mm, $t(25) = 4.41$, $p < 0.001$) listeners following sound offset. This decrease is also observed upon the subsequent sound exposure ($\text{change}_{\text{MISO}} = 0.03$ mm, $t(25) = 2.48$, $p = 0.02$; $\text{change}_{\text{NON}} = 0.04$ mm, $t(25) = 4.41$, $p < 0.001$), and occurred for some individual sound categories (see Supplemental Table 9).

Comparing the change in average pupil diameter during the final second of each sound to that of one second following sound offset

When controlling for sound duration, we observed that the pupil diameter did *not* reliably decrease toward baseline, when collapsed across all sounds, during first exposure for misophonic ($\text{change}_{\text{NON}} = 0.02$ mm, $t(25) = 1.55$, $p = 0.13$) and non-misophonic listeners ($\text{change}_{\text{MISO}} = 0.02$ mm, $t(25) = 1.62$, $p = 0.12$). However, upon second exposure, both misophonic and non-misophonic listeners showed a reliable decrease in pupil diameter after sound offset ($\text{change}_{\text{MISO}} = 0.02$ mm, $t(25) = 2.05$, $p = 0.05$; $\text{change}_{\text{NON}} = 0.03$ mm, $t(25) = 2.45$, $p = 0.02$). Similar trends were observed for some individual sound categories (see Supplemental Table 10).

Changes in average pupil diameter increases with reported emotion intensity

Thus far, our analyses have independently revealed that both reported emotion intensity and changes in pupil diameter are reliably greater for misophonic listeners toward sounds that evoke high levels of disgust and sounds that are common triggers. Despite the incomplete control of arousal across our sounds, focusing on the relationship between emotion intensity and pupil diameter allows us to better isolate the differences in emotion response between misophonic and non-misophonic listeners. We initially hypothesized a linear relationship between these two variables for both populations. For example, given that misophonic listeners report greater emotion intensity, their greater pupil dilation to certain sounds could be entirely a result of their emotional response. In that case, the misophonic linear function would be the same as non-misophonic listeners, but the misophonic responses would cluster in the upper right part of the function. Alternatively, it is possible that misophonics have greater pupillary reactivity for a given level of self-reported emotion intensity, resulting in a linear function positioned above that of non-misophonic listeners. To elucidate the type of relationship that exists between these two variables, we constructed a simple linear regression model using the *lm* function in the 'stat' package in R. Our model predicted the average change in pupil diameter using predictor variables of exposure (first or second), misophonic status (misophonic or non-misophonic), reported emotion intensity rating on a given scale, and the type of emotion scale (e.g., scale assessing intensity of disgust). We selected the change in pupil diameter at 2-to-3 seconds as our dependent variable because it was the most reliable measure of the physiological differences between misophonic and non-misophonic listeners for our sounds of interest (i.e., sounds in the *a priori* disgust and triggers category) (see*section B.A*).

We observed main effects of misophonic status ($F(1, 960) = 19.72, p < 0.001$), type of emotion scale ($F(9, 960) = 3.87$, $p < 0.001$), and exposure ($F(1, 960) = 13.40$, $p < 0.001$). There was a significant two-way interaction between emotion intensity rating **X** type of emotion scale ($F(9, 960) = 4.77$, $p < 0.001$), indicating that the correlation between intensity of emotions and change in pupil diameter depended upon the emotion scale. Furthermore, we observed a two-way interaction between exposure **X** misophonic status ($F(1, 960) = 5.72$, $p = 0.02$) which showed that the difference in pupil diameter between misophonic and non-misophonic listeners decreased upon second exposure. We did not observe any other interactions.

Figure 4 illustrates that the average change in pupil diameter scales linearly with emotion intensity for *many* emotion intensity scales. For both populations, no reliable association was observed for *excitement* (arousal) (p $<$ 0.5) or *fear* (p $<$ 0.3). Additionally, two scales exhibited contrasting patterns between our populations: 1) increases in *calmness* (arousal) reliably increased with pupil diameter for misophonic listeners, but not for non-misophonic listeners, while 2) increases in *sadness* did not reliably increase with pupil diameter for misophonic listeners, but it did for non-misophonic listeners. Importantly, misophonic listeners consistently exhibited greater pupil responses (dashed line) than non-misophonic listeners (solid line) in the scatterplots of pupil diameter versus emotion intensity A majority of these trends remained during second exposure to the sounds (see Supplemental Figure 2).

Figure 4. The average change in pupil diameter (mm) relative to baseline, at 2-to-3 seconds, across emotion intensity rating. Each panel represents a single emotion rating scale (e.g., *disgust*, *happiness*). Responses from misophonic listeners are represented in red, while responses from non-misophonic listeners are represented in blue. The solid, red line corresponds to the line-of-best fit across the data for misophonic listeners, while the dashed, blue line corresponds to the line-of-best fit across the data for non-misophonic listeners. The shaded region around the line-of-best fit depicts the 95% confidence interval. The data points represent individual sounds, whereby each unique symbol denotes the *a priori* emotion category the sound belongs to. A positive slope indicates that reported emotion intensity is associated with a greater average increase in pupil diameter from baseline.

To identify scales in which misophonic listeners exhibit greater pupil diameter changes than non-misophonic listeners despite reporting similar emotion intensity, we conducted a contrast analysis on the scales that showed a significant association between these two variables. We analzyed emotion intensity ratings at a medium intensity ('5') to limit data extrapolation. For negative emotions, misophonic listeners had a significantly greater change in pupil diameter, compared to non-misophonic listeners, when feeling *disgust* at medium intensity (t = -2.51, $p = 0.01$), feeling *annoyance* at medium intensity (t = -2.98, $p = 0.003$), and when perceiving the sound to have a medium intensity of *negative* valence (t = -2.30, $p = 0.02$). For positive emotions, misophonic listeners had reliably greater change in pupil diameter, compared to non-misophonic listeners, only when feeling *happiness* at a medium intensity (t = -2.08, $p = 0.04$).

IV. Discussion

Our study investigated whether there were measurable emotion intensity ratings, and pupillary differences for auditorily disgusting sounds compared to other emotional sounds between misophonics and non-misophonics. First, we found that misophonics report greater intensity of emotions when listening to sounds that commonly evoke disgust. In fact, they report feeling greater intensity of disgust, anger, annoyance, and fear at the time of listening to these sounds. Notably, misophonics felt these same emotions at similar intensities when listening to triggering, non-orofacial sounds (i.e., 'triggers') such as utensils scraping against one another, and water dripping from a faucet, while non-misophonics did not. Despite misophonic listeners anecdotally reporting that listening to these sounds produces a desire to escape the sound, we did not observe that misophonic listeners had a greater desire to escape the sounds mentioned previously, or any other category of sound. This outcome is not what we predicted given anecdotal reports from misophonia suffers of wanting to flee environments where trigger sounds are present, as well as items on clinical diagnostic scales that measure such behavior. However, a limitation of our study is that listening to sounds in an experimental setting is different from listening to sounds in the real world. Furthermore, the types of sounds in an experimental setting are not perfect exemplars of the sounds that result in the largest reactions because the sounds are not customized to the particular misophonic person's triggers, nor are they in a natural context where these sounds are typically heard. In general, our findings may have meaningful implications for future methodological designs that inform and describe clinical presentations of misophonia. Many questionnaires assessing severity of misophonia include questions about feelings of disgust and anger. However, most empirical studies do not ask about feelings of disgust and anger, but instead collect responses about other emotions, such as how sounds evoke anxiety, and intensity of bodily responses (e.g., Hansen et al., 2021; Edelstein et al., 2020; Samermit, Saal and Davidenko, 2019), as well as perceptions of the pleasantness of sounds (e.g., Enzler et al., 2021; Hansen et al., 2021). To our knowledge, our study is only one of two that has empirically measured and collected intensity responses of disgust and anger. Savard and colleagues (2022) were the first to collect ratings of anxiety, anger and disgust in response to trigger sounds changing within a multi-babble talker environment. They observed that the intensity of these emotions increased as the sounds became more identifiable to the listener. However, our study further contributes an understanding of which underlying emotions are distinct for misophonics, in general, and for misophonic trigger sounds in particular. Misophonics feel more intense emotions to many sounds which dilates their pupil diameter. They also dilate their pupils more even at the same level of emotion intensity of non-misophonic listeners. Given that we found measurable differences between our populations in the intensity of these emotions, we suggest that future research using misophonic populations include these scales to more accurately characterize emotion states in response to trigger sounds.

Second, we observed that misophonic listeners overall show greater pupillary reactivity to sounds that evoke high levels of disgust. This heightened physiological reaction is characterized by two features: (1) greater maximum pupil dilation, a difference of approximately 0.11 mm from non-misophonics, and (2) longer time course to peak arousal, a difference of approximately 630 ms from non-misophonics. The upward deflection we observe, suggestive of the pupil dilating relative to baseline, is a signature that has been found in other studies using emotional sounds (Nakakoga et al., 2020; Widmann, Schroger and Wetzel, 2018; Gingras et al., 2015; Partala and Surakka, 2003). Note, we did not observe a maximum pupil response at the offset of the sound as previously described in the literature (e.g., Winn et al., 2018; see in Gingras et al., 2015). Rather, we observed that the change in pupil diameter reliably decreased after maximum pupil dilation. There was no evidence that a second, maximum pupillary event occurred following stimulus offset. Our experimental paradigm included a one second delay period between the end of the sound and the presentation of the ratings. Thus, we did not have a large enough time window to assess how long the response sustains between our populations. It is possible that it would take much longer for misophonic listeners to return to a physiological baseline given the greater change in physiological reaction to the sound. We would encourage exploring the seconds after sound offset in a future study.

Our study is the first to investigate and report changes in pupil size to everyday sounds, including common misophonic triggers, in misophonic and non-misophonic listeners. To identify physiological differences across the populations, recent studies on misophonia have utilized functional magnetic resonance imaging (fMRI) to investigate resting state, and activity of regions during sound exposure (Neacsiu et al., 2022). At rest, individuals with mild presentations of misophonia show greater functional connectivity between the auditory and motor cortices (Kumar et al., 2021). Upon sound exposure, individuals with clinical presentations of misophonia have been found to show a greater amount of activity within the anterior insular cortex (AIC)11The AIC is known to be active when processing emotions, particularly disgust, anger and anxiety (Uddin et al., 2018). as well as increased functional connectivity between the AIC and regions comprising the default mode network (e.g., amygdala) (Kumar et al., 2017; Schroder et al., 2019; Hansen et al., 2022). In principle, a hyperactive AIC can produce a heightened sympathetic arousal response, given its projections to hypothalamic regions, which can phenotypically present as a larger increase in pupil diameter than typical.

Third, we observed a significant, positive association between the change in pupil diameter from baseline and emotional intensity rating across most emotion scales. Thus, reporting a greater intensity of emotion indicates that an individual is having a physiological reaction congruent in magnitude. This was true for both misophonic and non-misophonic listeners, with misophonic responses overall being larger than non-misophonic responses. Our findings are consistent with previous research that has shown greater physiological reactions for sounds that evoke greater intensity of emotions (Partala and Surakka, 2003; Bradley and Lang, 2000). There are some studies that find an equal pupillary response for sounds of different valence (e.g., Partala and Surakka, 2003). Our selected sounds only come from a small corpus of the possible everyday sounds; therefore, it was not possible to properly analyze sounds differing in valence while controlling for the same intensity of arousal.

Lastly, we observed that emotion intensity and pupillary reaction toward sounds remain consistent upon second exposure. Thus, under our time scale, there appears to be no evidence of increased responses (such as might be caused by the mere exposure effect) or decreased responses (such as might be caused by habituation). Our findings provide support for Schroder et al. (2019) in that repeated exposure to trigger sounds does not diminish the emotional or physiological response that occurs.

V. Conclusion

By investigating physiological responses through pupillometry in misophonic listeners, our study contributes a set of findings to the ongoing work on understanding the clinical presentations of misophonia. First, we observed that misophonic listeners exhibit heightened emotional intensity, reporting increased levels of disgust, anger, annoyance, and fear when exposed to sounds that universally evoke high levels of disgust. This same difference between populations occurs for non-orofacial trigger sounds; this is a unique aspect of misophonia because these sounds are not perceived nor evoke feelings of disgust in non-misophonic populations. Second, we observed that misophonic listeners exhibit greater pupillary reactivity to highly disgusting sounds, characterized by both a larger maximum pupil dilation and a longer time course to peak arousal compared to non-misophonic listeners. These findings suggest a heightened sympathetic arousal response in individuals with misophonia, potentially rooted in hyperactivity of the anterior insular cortex. Third, and most notably, we observed a significant correlation between reported emotional intensity and pupillary responses, indicating that physiological arousal via changes in pupil diameter can be non-invasively used to measured emotional responses in the absence of self-reported emotion, particularly towards sounds evoking disgust, anger, and annoyance. When used in tandem, integrating subjective assessments with objective physiological measures like pupillometry can be a powerful way of gauging treatment effectiveness in a clinical setting.

Author contributions

U. Oszczapinska : Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Resources; Supervision; Validation; Visualization; Writing - original draft; Writing - review & editing. **S. Park** : Data curation; Formal analysis; Investigation; Software; Validation; Visualization; Writing - original draft; Writing - review & editing. **Y. Qiu** : Data curation; Formal analysis; Software; Validation. **B. Nance** : Conceptualization; Investigation; Methodology; Resources. **M. Julien** : Conceptualization; Investigation; Methodology; Resources.**L. M. Heller** *:* Conceptualization; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Writing - review & editing

Acknowledgments

This research was supported and funded by the Misophonia Research Fund (MRF). Open Access funding awarded by SoQuiet.

jabbrv-ltwa-all.ldf jabbrv-ltwa-en.ldf**Conflict of interest statement**

The authors declare no competing interests.

References

Acoustics-Assessment of Noise Annoyance by Means of Social and Socio-Acoustic

Surveys; 2021. www.iso.org

American speech-language-hearing association. (2024). American Speech-Language-Hearing Association. https://www.asha.org/

Ash, P. A., Benzaquén, E., Gander, P. E., Berger, J. I., & Kumar, S. (2024). Mimicry in misophonia: A large-scale survey of prevalence and relationship with trigger sounds. *Journal of clinical psychology* ,*80* (1), 186-197.

Ayzenberg, V., Hickey, M. R., & Lourenco, S. F. (2018). Pupillometry reveals the physiological underpinnings of the aversion to holes. *PeerJ* , *6* , e4185.

Ballas, J. A., & Howard Jr, J. H. (1987). Interpreting the language of environmental sounds. *Environment and behavior* , *19* (1), 91-114.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of experimental psychology: human perception and performance* , *19* (2), 250.

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48.

Birren, J. E., Casperson, R. C., & Botwinick, J. (1950). Age changes in pupil size. *Journal of Gerontology* , *5* (3), 216-221.

Bitsios, P., Prettyman, R., & Szabadi, E. (1996). Changes in autonomic function with age: a study of pupillary kinetics in healthy young and old people. *Age and ageing* , *25* (6), 432-438.

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology* , *37* (2), 204–215.

Bradley, M. M., & Lang, P. J. (2007). *The International Affective Digitized Sounds (; IADS-2): Affective ratings of sounds and instruction manual* (Vol. 434). Technical report B-3. University of Florida, Gainesville, Fl.

Brickman, P., Redfield, J., Harrison, A. A., Crandall, R. (1972). Drive and predisposition as factors in the attitudinal effects of mere exposure. *Journal of Experimental Social Psychology* , 8, 31–44.

Cox, T. J. (2008). Scraping sounds and disgusting noises. *Applied Acoustics* , *69* (12), 1195-1204.

Craig, D. G. (2005). An exploratory study of physiological changes during "chills" induced by music. *Musicae scientiae* ,*9* (2), 273-287.

Edelstein, M., Brang, D., Rouw, R., & Ramachandran, V. S. (2013). Misophonia: physiological investigations and case descriptions.*Frontiers in Human Neuroscience* , *7* , 47988.

Edelstein, M., Monk, B., Ramachandran, V. S., & Rouw, R. (2020). Context influences how individuals with misophonia respond to sounds.*biorxiv* , 2020-09.

Enzler, F., Loriot, C., Fournier, P., & Norena, A. J. (2021). A psychoacoustic test for misophonia assessment. *Scientific Reports* , *11* (1), 11044.

Freed, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *The Journal of the Acoustical Society of America* , *87* (1), 311-322.

Gabrielsson, A., & Juslin, P. N. (2003). *Emotional expression in music* . Oxford University Press.

Goldstein, A. (1980). Thrills in response to music and other stimuli.*Physiological psychology* , *8* (1), 126-129.

Grossini, E., Stecco, A., Gramaglia, C., De Zanet, D., Cantello, R., Gori, B., ... & Zeppegno, P. (2022). Misophonia: Analysis of the neuroanatomic patterns at the basis of psychiatric symptoms and changes of the orthosympathetic/parasympathetic balance. *Frontiers in Neuroscience* , *16* , 827998.

Grassi, M. (2005). Do we hear size or sound? Balls dropped on plates.*Perception & psychophysics* , *67* , 274-284.

Gingras, B., Marin, M. M., Puig-Waldmuller, E., & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in human neuroscience* ,*9* , 619.

Halpern, D. L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception & Psychophysics* , *39* , 77-80.

Hansen, H. A., Leber, A. B., & Saygin, Z. M. (2021). What sound sources trigger misophonia? Not just chewing and breathing. *Journal of Clinical Psychology* , *77* (11), 2609-2625.

Hansen, H. A., Stefancin, P., Leber, A. B., & Saygin, Z. M. (2022). Neural evidence for non-orofacial triggers in mild misophonia.*Frontiers in Neuroscience* , *16* , 880759.

Heller, L. M., & Smith, J. M. (2022). Identification of everyday sounds affects their pleasantness. *Frontiers in Psychology* , *13* , 894034.

Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. Behavior Research Methods, 50(1), 107–114.

Hevner, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology* , *47* (1), 103-118.

Houben, M. M. J. (2002). The sound of rolling objects: perception of size and speed.

Hunter, P. G., & Schellenberg, E. G. (2010). Music and emotion.*Music perception* , 129-164.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception* , *23* (4), 319-330.

Jastreboff, M. M., & Jastreboff, P. J. (2001). Components of decreased sound tolerance: hyperacusis, misophonia, phonophobia. *ITHS News Lett* , *2* (5-7), 1-5.

Jastreboff, M. M., & Jastreboff, P. J. (2002). Decreased Sound Tolerance and Tinnitus Retraining Therapy (TRT). The Australian and New Zealand Journal of Audiology, 24(2), 74–84.

Jastreboff, P. J., & Jastreboff, M. M. (2023). The neurophysiological approach to misophonia: Theory and treatment. *Frontiers in Neuroscience* , *17* , 895574.

Kazazis, S., Korsmit, I., & McAdams, S. (2024). Triggering Misophonia: The Importance of Spectral Information, Temporal Information, and Action Identification. *Auditory Perception & Cognition* , 1-13.

Khalfa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music?. *International Journal of Psychophysiology* , *68* (1), 17-26.

Klatzky, R. L., Pai, D. K., & Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence* , *9* (4), 399-410.

Kumar, S., Tansley-Hancock, O., Sedley, W., Winston, J. S., Callaghan, M. F., Allen, M., . . . & Griffiths, T. D. (2017). The brain basis for misophonia. *Current Biology* , *27* (4), 527-533.

Kumar, S., Dheerendra, P., Erfanian, M., Benzaquen, E., Sedley, W., Gander, P. E., . . . & Griffiths, T. D. (2021). The motor basis for misophonia. *Journal of Neuroscience* , *41* (26), 5762-5770.

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.

Lee, D. H., & Anderson, A. K. (2017). Reading What the Mind Thinks From How the Eye Sees. *Psychological Science* , *28* (4), 494-503.

Lemaitre, G., & Heller, L. M. (2012). Auditory perception of material is fragile while action is strikingly robust. *The Journal of the Acoustical Society of America* , *131* (2), 1337-1348.

Lenth, R. V. (2024). emmeans: Estimated Marginal Means, aka Least-Squares Means (1.10.0) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lobato-Rincon, L. L., Cabanillas-Campos, M.delC., Bonnin-Arias, C., Chamorro-Gutierrez, E., Murciano-Cespedosa, A., & Sanchez-Ramos Roda, C. (2014). Pupillary behavior in relation to wavelength and age. *Frontiers in human neuroscience* , *8* , 221.

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.

Nakakoga, S., Higashi, H., Muramatsu, J., Nakauchi, S., & Minami, T. (2020). Asymmetrical characteristics of emotional responses to pictures and sounds: Evidence from pupillometry. *PloS one* , *15* (4), e0230775.

Nardelli, M., Valenza, G., Greco, A., Lanata, A., & Scilingo, E. P. (2015). Recognizing emotions induced by affective sounds through heart rate variability. *IEEE Transactions on Affective Computing* ,*6* (4), 385-394.

Neacsiu, A. D., Szymkiewicz, V., Galla, J. T., Li, B., Kulkarni, Y., & Spector, C. W. (2022). The neurobiology of misophonia and implications for novel, neuroscience-driven interventions. *Frontiers in Neuroscience* , *16* , 893903.

The MathWorks Inc. (2020b). MATLAB version: 9.13.0 (R2020b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

The MathWorks Inc. (2022a). MATLAB version: 9.13.0 (R2022a), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

Oszczapinska, U., Heller, L. M., Jang, S., & Nance, B. (2024). Ecological sound loudness in environmental sound representations.*JASA Express Letters* , *4* (2).

Panksepp, J. (1995). The emotional sources of" chills" induced by music.*Music perception* , *13* (2), 171-207.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International journal of human-computer studies* , *59* (1-2), 185-198.

Piczak, K. J. (2015, September). Environmental sound classification with convolutional neural networks. In *2015 IEEE 25th international workshop on machine learning for signal processing (MLSP)* (pp. 1-6). IEEE.

Radsten Ekman, M., Lunden, P., & Nilsson, M. E. (2015). Similarity and pleasantness assessments of water-fountain sounds recorded in urban public spaces. *The Journal of the Acoustical Society of America* ,*138* (5), 3043-3052.

R Core Team. (2024). R: A Language and Environment for Statistical Computing (4.3.3) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rosenthal, M. Z., Anand, D., Cassiello-Robbins, C., Williams, Z. J., Guetta, R. E., Trumbull, J., & Kelley, L. D. (2021). Development and initial validation of the duke misophonia questionnaire. *Frontiers in psychology* , *12* , 709928.

Samermit, P., Saal, J., & Davidenko, N. (2019). Cross-sensory stimuli modulate reactions to aversive sounds. *Multisensory Research* ,*32* (3), 197-213.

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* , *44* (2), 293-304.

Savard, M. A., Sares, A. G., Coffey, E. B., & Deroche, M. L. (2022). Specificity of affective responses in misophonia depends on trigger identification. *Frontiers in Neuroscience* , *16* , 879583.

Schroder, A., Vulink, N., & Denys, D. (2013). Misophonia: diagnostic criteria for a new psychiatric disorder. *PloS one* , *8* (1), e54706.

Schroder, A., van Wingen, G., Eijsker, N., San Giorgi, R., Vulink, N. C., Turbyne, C., & Denys, D. (2019). Misophonia is associated with altered brain activity in the auditory cortex and salience network.*Scientific reports* , *9* (1), 7542.

Sherlock, L. P., & Formby, C. (2005). Estimates of loudness, loudness discomfort, and the auditory dynamic range: normative estimates, comparison of procedures, and test-retest reliability. *Journal of the American Academy of Audiology* , *16* (02), 085-100.

Shulman, A., & Goldstein, B. (1996). A Final Common Pathway for Tinnitus-Implications for Treatment. *The International Tinnitus Journal* , *2* , 137-142.

Siepsiak, M., Vrana, S. R., Rynkiewicz, A., Rosenthal, M. Z., & Dragan, W. Ł. (2023). Does context matter in misophonia? A multi-method experimental investigation. *Frontiers in neuroscience* , *16* , 880853.

Swedo, S. E., Baguley, D. M., Denys, D., Dixon, L. J., Erfanian, M., Fioretti, A., . . . & Raver, S. M. (2022). Consensus definition of misophonia: a delphi study. *Frontiers in Neuroscience* , *16* , 841816.

Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and Function of the Human Insula. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society* , *34* (4), 300–306.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition & Emotion* , *22* (4), 720-752.

Vitoratou, S., Hayes, C., Uglik-Marucha, N., Pearson, O., Graham, T., & Gregory, J. (2023). Misophonia in the UK: Prevalence and norms from the S-Five in a UK representative sample. *Plos one* , *18* (3), e0282777.

Warren, W. H., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: a case study in ecological acoustics.*Journal of Experimental Psychology: Human perception and performance* , *10* (5), 704.

Widmann, A., Schröger, E., & Wetzel, N. (2018). Emotion lies in the eye of the listener: Emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3.*Biological psychology* , *133* , 10-17.

Williams, Z. J., Cascio, C. J., & Woynaroski, T. G. (2022). Psychometric validation of a brief self-report measure of misophonia symptoms and functional impairment: The duke-vanderbilt misophonia screening questionnaire. *Frontiers in Psychology* , *13* , 897901.

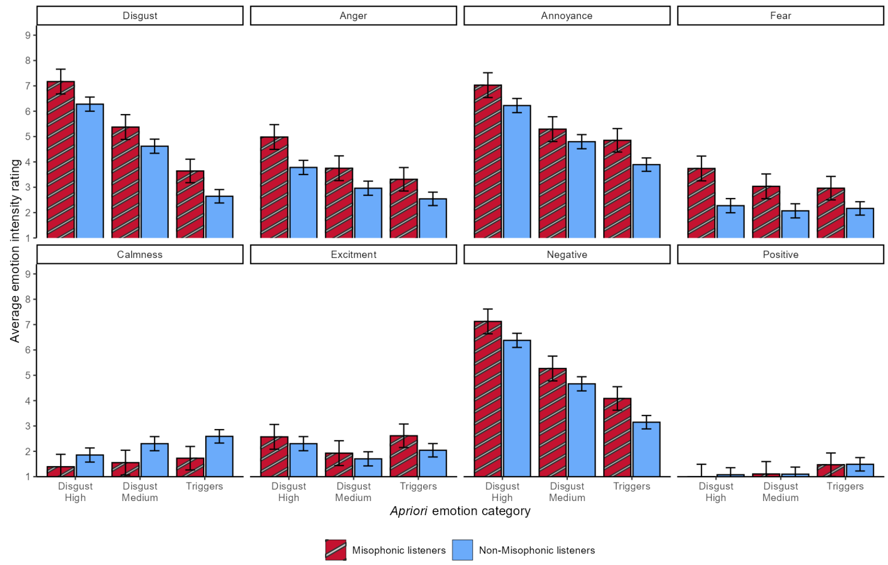
Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects.*Investigative ophthalmology & visual science* , *35* (3), 1132-1137.

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in hearing* , *22* , 2331216518800869.

Witvliet, C. V., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion* , *21* (1), 3-25.

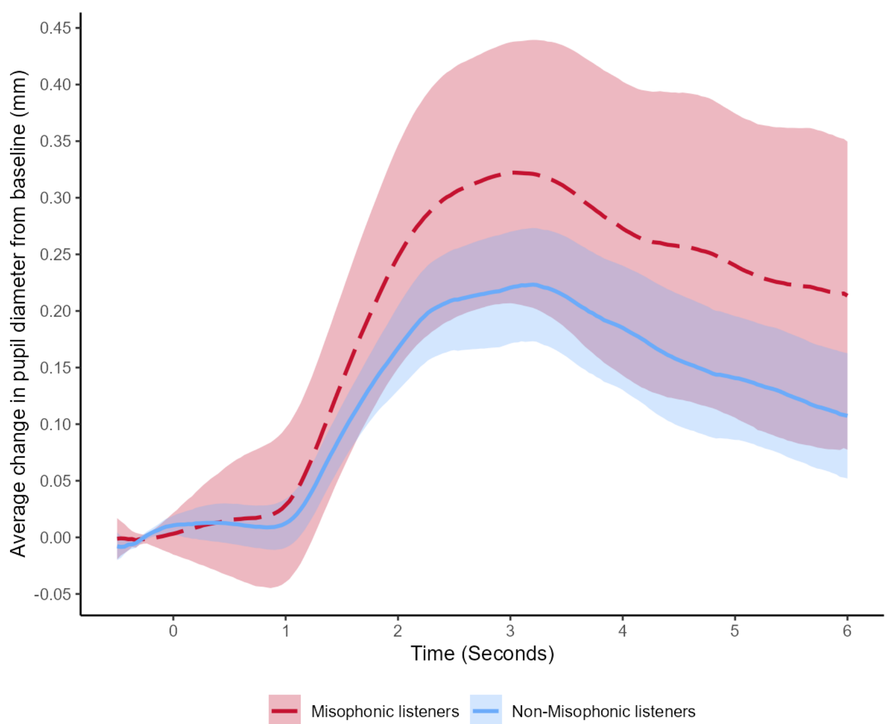
Wu, M. S., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2014). Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. *Journal of clinical psychology* ,*70* (10), 994-1007.

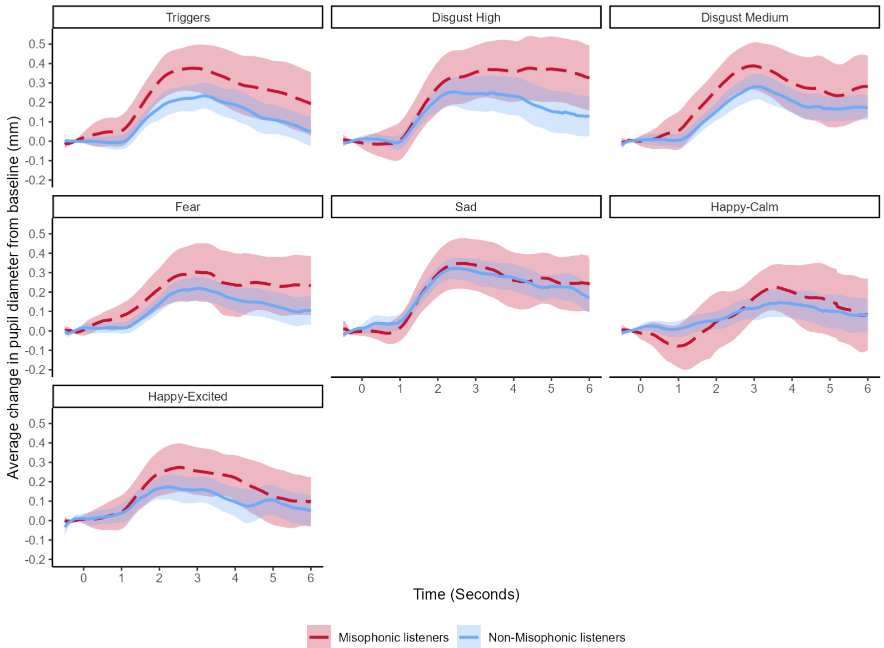
Zajonc, R. B. (1968). Attitudinal effects of mere exposure.*Journal of personality and social psychology* , *9* (2p2), 1.

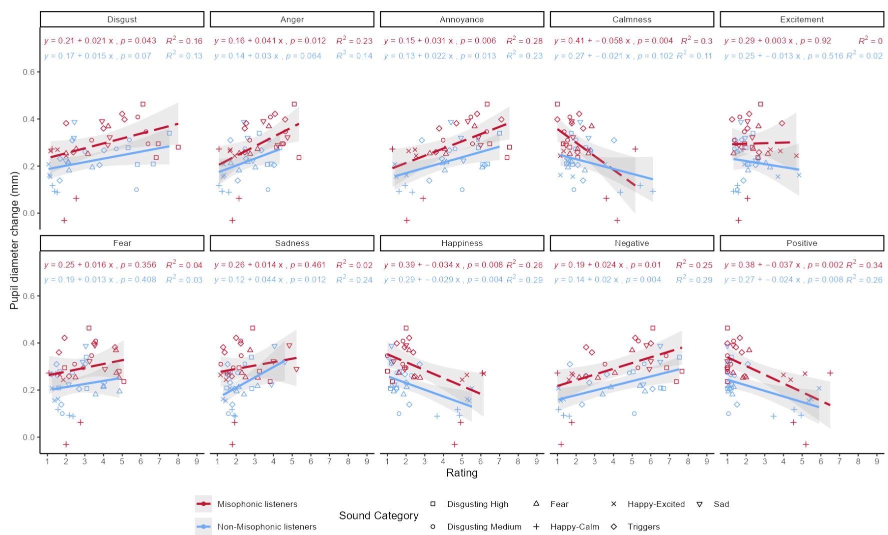


Hosted file

`Table 1.docx` available at https://authorea.com/users/845311/articles/1233906-the-impact-of-disgusting-sounds-on-pupil-diameter-of-misophonic-and-non-misophonic-listeners




Triggers
Disgust High
Disgust Medium
Fear
Sad
Happy-Calm
Happy-Excited
Average change in pupil diameter from baseline (mm)
Time (Seconds)
Misophonic listeners
Non-Misophonic listeners


Disgust
$y = 0.21 + 0.021\,x$, $p = 0.043$ $R^2 = 0.16$
$y = 0.17 + 0.015\,x$, $p = 0.07$ $R^2 = 0.13$
Anger
$y = 0.16 + 0.041\,x$, $p = 0.012$ $R^2 = 0.23$
$y = 0.14 + 0.03\,x$, $p = 0.064$ $R^2 = 0.14$
Annoyance
$y = 0.15 + 0.031\,x$, $p = 0.006$ $R^2 = 0.28$
$y = 0.13 + 0.022\,x$, $p = 0.013$ $R^2 = 0.23$
Calmness
$y = 0.41 + -0.058\,x$, $p = 0.004$ $R^2 = 0.3$
$y = 0.27 + -0.021\,x$, $p = 0.102$ $R^2 = 0.11$
Excitement
$y = 0.29 + 0.003\,x$, $p = 0.92$ $R^2 = 0$
$y = 0.25 + -0.013\,x$, $p = 0.516$ $R^2 = 0.02$
Fear
$y = 0.25 + 0.016\,x$, $p = 0.356$ $R^2 = 0.04$
$y = 0.19 + 0.013\,x$, $p = 0.408$ $R^2 = 0.03$
Sadness
$y = 0.26 + 0.014\,x$, $p = 0.461$ $R^2 = 0.02$
$y = 0.12 + 0.044\,x$, $p = 0.012$ $R^2 = 0.24$
Happiness
$y = 0.39 + -0.034\,x$, $p = 0.008$ $R^2 = 0.26$
$y = 0.29 + -0.029\,x$, $p = 0.004$ $R^2 = 0.29$
Negative
$y = 0.19 + 0.024\,x$, $p = 0.01$ $R^2 = 0.25$
$y = 0.14 + 0.02\,x$, $p = 0.004$ $R^2 = 0.29$
Positive
$y = 0.38 + -0.037\,x$, $p = 0.002$ $R^2 = 0.34$
$y = 0.27 + -0.024\,x$, $p = 0.008$ $R^2 = 0.26$
Pupil diameter change (mm)
Rating
Misophonic listeners
Non-Misophonic listeners
Sound Category
Disgusting High
Disgusting Medium
Fear
Happy-Calm
Happy-Excited
Triggers
Sad