BRILL Multisensory Research (2024) DOI:10.1163/22134808-bja10137 brill.com/msr

Positive Attributable Visual Sources Attenuate the Impact of Trigger Sounds in Misophonia

Ghazaleh Mahzouni *, **Moorea M. Welch**, **Michael Young**, **Veda Reddy**, **Patrawat Samermit and Nicolas Davidenko**

Psychology Department, University of California, Santa Cruz, CA 95064, USA
*Corresponding author; e-mail: gmahzoun@ucsc.edu
ORCID iDs: Mahzouni: 0009-0004-2619-4269; Samermit: 0009-0001-2069-9416; Davidenko: 0000-0002-6935-5542

Received 21 November 2023; accepted 4 November 2024; published online 12 November 2024

Abstract
Misophonia is characterized by strong negative reactions to everyday sounds, such as chewing, slurping or breathing, that can have negative consequences for daily life. Here, we investigated the role of visual stimuli in modulating misophonic reactions. We recruited 26 misophonics and 31 healthy controls and presented them with 26 sound-swapped videos: 13 trigger sounds paired with the 13 Original Video Sources (OVS) and with 13 Positive Attributable Visual Sources (PAVS). Our results show that PAVS stimuli significantly increase the pleasantness and reduce the intensity of bodily sensations associated with trigger sounds in both the misophonia and control groups. Importantly, people with misophonia experienced a larger reduction of bodily sensations compared to the control participants. An analysis of self-reported bodily sensation descriptions revealed that PAVS-paired sounds led participants to use significantly fewer words pertaining to body parts compared to the OVS-paired sounds. We also found that participants who scored higher on the Duke Misophonia Questionnaire (DMQ) symptom severity scale had higher auditory imagery scores, yet visual imagery was not associated with the DMQ. Overall, our results show that the negative impact of misophonic trigger sounds can be attenuated by presenting them alongside PAVSs.

Keywords
attributable visual source, bodily sensations, misophonia, multisensory perception, sound-swapped videos, trigger sounds

Published with license by Koninklijke Brill BV | DOI: 10.1163/22134808-bja10137
 | ISSN: 2213-4794 (print) 2213-4808 (online)

1. Introduction

For some individuals, seemingly ordinary everyday sounds can impose disproportionate levels of distress and discomfort in their lives. Misophonia is a relatively new disorder that describes this extreme psychological and physiological reaction to specific types of trigger sounds and associated stimuli (Jastreboff and Jastreboff, 2001; Swedo *et al.*, 2022). Common misophonia triggers include human-generated eating sounds (e.g., chewing, slurping, crunching, lip smacking), nasal sounds (snoring, coughing, sniffling), and various repetitive sounds such as pen tapping, refrigerator humming, clock ticking, etc. (Schröder *et al.*, 2013; Jager *et al.*, 2020). An important characteristic of the misophonic experience is the inordinate aversive response to the trigger sound, which can compromise day-to-day activities, school or work life, and healthy social interactions (Edelstein *et al.*, 2013). Trigger sounds often elicit an intense and automatic negative emotional response (e.g., distress, anger, and/or disgust toward the person generating the sound) and sympathetic (fight or flight) physiological response (e.g., increased heart rate, bodily tension) in people with misophonia (Edelstein *et al.*, 2013; Jager *et al.*, 2020; Kumar *et al.*, 2017; Rouw and Erfanian, 2018).

In a recent large-scale study with a representative sample of the US population, Dixon *et al.* (2024) found that 78.5% of the sample reported sensitivity to misophonia trigger sounds, and importantly 4.6% of individuals in the US reported clinical levels of misophonia symptoms which were accompanied by severe functional impairment in social and private activities, home management and the ability to maintain close relationships. Despite the prevalence, misophonia is not yet an officially recognized disorder in the DSM-V, so there are neither established clinical diagnostic criteria for misophonia nor established treatments. In most research studies, misophonia is assessed by self-report measures of misophonia symptoms through questionnaires such as the Misophonia Questionnaire (MQ; Wu *et al.*, 2014). Amsterdam- Misophonia Scale (A-MISO-S; Schröder *et al.*, 2013), and most recently, the S-Five (Vitoratou *et al.*, 2021) and the Duke Misophonia Questionnaire scale (DMQ; Rosenthal *et al.*, 2021).

A recent neuroimaging investigation has suggested that trigger sounds are, in fact, processed differently in the brains of people with misophonia compared to other aversive sounds. Kumar *et al.* (2017) asked participants to listen and rate misophonic trigger sounds (e.g., chewing), universally unpleasant sounds (e.g., screaming), or neutral sounds (e.g., rain) and measured brain activity using fMRI. As expected, those who reported having misophonia rated the trigger sounds as more annoying compared to unpleasant or neutral sounds, whereas those without misophonia rated the universally unpleasant sounds as the most annoying. Critically, for the misophonic group, the trigger sounds evoked a larger response in the anterior insular cortex (AIC), as well as functional hyperconnectivity between areas of the brain that are responsible for emotional regulation, emotional behavior, and memories of past events, including the frontal cortex, amygdala, and hippocampus. The fact that these differences are in higher-level brain regions, as opposed to low-level auditory regions, suggests that misophonic responses may be modulated by high-level cognitive factors and multisensory processes.

The same study also showed that people with misophonia show an increased heart rate and galvanic skin response (GSR) in response to hearing trigger sounds (e.g., chewing) but not generally unpleasant (e.g., baby screaming) or neutral sounds (e.g., rainfall) in comparison to those without misophonia (Kumar *et al.*, 2017; also see Grossini *et al.*, 2022; Neacsiu *et al.*, 2023). Additionally, subjective reports of body perception as measured by the body consciousness questionnaire showed that misophonia participants had a greater tendency to focus on internal bodily sensations compared to healthy controls. Kumar *et al.* (2017) argue that this finding, coupled with the abnormal functioning of AIC, contributes to enhanced saliency of trigger sounds for misophonia participants.

Initially, misophonia was regarded as an audiological disorder comparable to decreased sound tolerance disorders such as hyperacusis. However, recent findings suggest that, unlike hyperacusis, reactions to misophonia trigger sounds are not tied to specific physical characteristics of the sounds, such as pitch, volume, or timbre (Jastreboff and Jastreboff, 2014), and context may play an important role in how a trigger sound is experienced. For example, Savard *et al.* (2022) embedded varying levels of noise in misophonia trigger sounds to make them less identifiable and asked participants to rate their subjective emotional reaction to the sound. They found that when trigger sounds were easily identifiable, participants with higher misophonia scores were more likely to report negative emotional reactions to the sounds than when the sounds were difficult to identify, which suggests a role of higher-level cognition in interpreting and reacting to trigger sounds. Similarly, Heller and Smith (2022) showed that misidentifying the source of a trigger sound as belonging to a more neutral sound category significantly increased the perceived pleasantness of the sound.

The multisensory aspect of the misophonic experience has received less attention. Previous research has established that varying visual contexts can modulate reactions to aversive sounds in the general population. For example, Samermit *et al.* (2019) found that presenting a Positive Attributable Visual Source (PAVS) of an aversive sound reduced negative reactions to that sound in a neurotypical sample. In their study, participants rated aversive sounds presented either with the corresponding Original Video Source (OVS) of the sound (e.g., a fork scraping glass), with a PAVS (e.g., someone jumping on a bed), or with no visual source. The results showed that pairing aversive sounds with a PAVS significantly reduced negative reactions to the aversive sounds, including discomfort, unpleasantness, and bodily sensations, compared to presenting the sounds alone or with the OVS. In another study, Siepsiak *et al.* (2022) found that human smacking mouth sounds paired with incongruent visual stimuli (e.g., a video of a hand playing in mud) reduced self-reported negative reactions to the trigger sound compared to human or animal eating sounds presented with their corresponding congruent video stimuli.

Recently, Samermit *et al.* (2022) developed a new validated set of 36 *sound-swapped video* stimuli presenting common misophonia trigger sounds (e.g., crunchy chewing) paired with both OVS (e.g., someone biting into a chip) and PAVS (e.g., tearing a piece of paper; see http://osf.io/3ysfh/). In a study with 102 neurotypical participants, observers were asked to rate each trigger sound in terms of its pleasantness or unpleasantness in the PAVS or OVS context. Results showed that misophonia trigger sounds presented in the PAVS context were rated as significantly more pleasant than the same sounds presented in the OVS context. Although the study was conducted on a neurotypical population, the effectiveness of the PAVS *vs* OVS contexts was observed across individuals with varying levels of misophonia symptoms.

An open question remains about the effectiveness of PAVS stimuli for a wider range of trigger sounds in the misophonia population specifically. The present study aims to answer this question by examining a group with misophonia and a control group to measure whether the effectiveness of PAVS stimuli in attenuating the negative reactions to trigger sounds is similar or different between groups.

We hypothesized that the PAVS stimuli would increase pleasantness ratings of trigger sounds and decrease the intensity of bodily sensations elicited by trigger sounds in both groups. Moreover, we hypothesized that there would be a positive relationship between visual and auditory imagery abilities and how people experience PAVS *vs* OVS stimuli across pleasantness and bodily sensation scales. Previous work with neurotypical participants has shown that mental imagery relies on overlapping cognitive and brain mechanisms as actual perception and that mental imagery, in one sense, can modulate perception in a different sense (Berger and Ehrsson 2013, 2014; 2018; Kosslyn, 1995; Pearson *et al.*, 2008). For example, Berger and Ehrsson (2013) showed that simply imagining an auditory /ba/ speech sound while viewing silent /ga/ lip movements can lead to a classic McGurk illusion of perceiving /da/. Given that both McGurk stimuli and sound-swapped stimuli involve incongruent audiovisual stimuli where the visual input overrides auditory perception, we reasoned that multisensory imagery could help integrate the slightly mismatched auditory and visual inputs in the PAVS stimuli in a similar fashion. Although participants in our study were not asked to engage in mental imagery, we still predict that these processes, which may happen automatically, may aid participants in integrating the visual and auditory information and experiencing the benefits of PAVS stimuli.

Additionally, previous work has shown people with misophonia report having more clear and vivid auditory imagery compared to people without misophonia (Simner *et al.*, 2021), but differences in visual imagery have not been investigated. As such, we administered the Vividness of Visual Imagery Questionnaire (VVIQ), The Bucknell Auditory Imagery Scale for Vividness (BAIS-V), and the Clarity of Auditory Imagery Scale Vividness (CAIS) to all participants to investigate whether the capacity for visual and auditory mental imagery vividness might facilitate cross-sensory remapping.

2. Method

All participants were assessed on the three misophonia scales: MQ, AMISO-S, and DMQ. We have previously used MQ in our Samermit *et al.* (2022) study, and at the time of conducting the current study, DMQ was the most recent and comprehensive measure of misophonia. For completeness, we also included AMISO-S to see if all measures produced similar results. Given that all three measures were highly correlated (see Table 5 below), we focused our analysis on a composite scale derived from the combined affective, physiological, and cognitive subscales, which we refer to as the *DMQ symptom severity score* ranging from 0 to 4 (see Section 2.4. *Measures*).

2.1. Participants

We recruited 31 misophonic participants who were interviewed in a previous study (Samermit *et al.*, 2022). We excluded five participants who scored lower than 1.5 in the DMQ symptom severity scale, which is calculated based on the average responses on the affective, physiological, and cognitive subscales. As such, the misophonia group included a total of 26 participants (20 females, five males, and one other, ages 20 to 63, $M = 27.08$, SD = 8.73) who received a US \$40 Amazon gift card for their participation. We also recruited 38 naive participants from the psychology research pool at the University of California Santa Cruz (UCSC) as the control group. We excluded seven control participants who scored higher than 1.5 in the DMQ symptom severity scale; thus, the control group included a total of 31 participants (26 females, three males, and two others, ages 18 to 61; $M = 21.55$, SD = 7.80) who received course credit for their participation. Because participants in the misophonia group were recruited *via* ads from the general population, it includes a broader age range compared to the control group. An independent-sample t-test revealed that the average age for the misophonia group (27.08) was significantly higher than for the control group (21.55; $t_{55} = 2.52$, $p = 0.01$); however, we have

Table 1.
Self-report misophonia scores across three different misophonia questionnaires

Scale	Misophonia group ($n = 26$)		Control group ($n = 31$)	
	M (SD)	Cronbach's alpha	*M* (SD)	Cronbach's alpha
DMQ symptom scale (0 to 4)	2.40 (0.54)	0.84	0.82 (0.28)	0.91
A-MISO-S (0 to 4)	10.11 (3.01)	0.73	3.87 (2.50)	0.61
MQ sensitivity scale (0 to 4)	36.73 (8.29)	0.73	15.10 (10.79)	0.88
MQ severity scale (0 to 15)	5.58 (1.77)	–	2.71 (2.16)	–

Note. Means with standard deviations in parentheses for Duke misophonia questionnaires (DMQ) symptom scale, Misophonia questionnaire (MQ), and Amsterdam misophonia scale (A-MISO-S). Cronbach's alpha cannot be calculated for the MQ severity scale because this scale contains a single item.

no reason to believe this age difference should play a significant role in our results. This study was approved by the Institutional Review Board (IRB) at UCSC. Table 1 shows the self-reported DMQ symptom severity scores, A-MISO-S, MQ-Sensitivity, and MQ-Severity scores for the misophonia and control groups.

2.2. Stimuli

All stimuli were selected from the sound-swapped video database (https://osf.io/3ysfh/) published by Samermit *et al.* (2022). We selected the 13 most effective pairs of OVS and their corresponding PAVS stimuli, meaning that these stimuli produced the greatest difference in pleasantness scores (PAVS–OVS) across the general population. The 13 sounds span 10 common trigger categories: slurping, crunchy chewing, wet chewing, scraping, squeaking, sniffling, swishing, tapping, and clicking. Table 2 includes a list and description of all 13 PAVS and 13 OVS stimuli used in the current study.

2.3. Procedure

Participants completed the experiment *via* Qualtrics. After filling out a consent form, they proceeded to watch 13 PAVS stimuli in a random order and then the 13 OVS stimuli in a random order. We previously found that this order of presentation (PAVS before OVS) leads to a larger modulation of reactions compared to presenting the OVS stimuli first (Samermit *et al.*, 2022). After watching each video, participants were asked to provide ratings of how pleasant or unpleasant they found the sound and the intensity of any experienced bodily sensations. After rating all of the videos, participants filled out several questionnaires, including three misophonia questionnaires: the DMQ, Amsterdam Misophonia Scale (A-MISO-S), and MQ. Subsequently, participants

Table 2.
List of stimuli used in the experiment

Video number	Trigger sound category	OVS descriptions	PAVS descriptions
1 (Vid01_O/P)	Slurping	Actor sips empty drink through straw	Water flows through rock in creek
2 (Vid04_O/P)	Crunchy chewing	Actor eats a cheese crisp	Actor tears a piece of paper
3 (Vid09_O/P)	Scraping	Fork scraping on a plate	Turning a faucet
4 (Vid06_O/P)	Squeaking	Cleaning a window	Puppies barking
5 (Vid07_O/P)	Swishing	Swishing water in mouth	Shaking a water bottle
6 (Vid08_O/P)	Sniffling	Actor sniffling loudly	Sweeping floor with a broom
7 (Vid10_O/P)	Crunchy chewing	Actor crunching with closed mouth	Stick in mud
8 (Vid12_O/P)	Sniffling	Actor sniffling loudly	Pencil drawing on paper
9 (Vid13_O/P)	Tapping	Fork tapping on metal knife	Hammering on nail
10 (Vid16_O/P)	Drumming	Finger drumming on table	Popcorn popping
11 (Vid17_O/P)	Wet chewing	Actor chewing with mouth open	Draining a towel
12 (Vid18_O/P)	Gulping	Actor gulps water	Stirring stick in water
13 (Vid20_O/P)	Clicking	Actor clicks a pen	Bike wheel spinning

Note. The names in parentheses correspond to the names used in the Sound-Swapped Video database: https://osf.io/3ysfh/

completed the VVIQ, BAIS-V, and CAIS. Finally, participants answered a set of demographic questions.

2.4. Measures

2.4.1. Pleasantness and Bodily Sensation Ratings

Pleasantness ratings were reported on the following seven-point scale: 1, Extremely unpleasant; 2, Very unpleasant; 3. Somewhat unpleasant; 4, Neither pleasant nor unpleasant; 5, Somewhat pleasant; 6, Very pleasant; and 7, Extremely pleasant. Ratings of the intensity of bodily sensations were reported on the following seven-point scale: 1, No sensation; 2, 3, Mild sensation; 4, 5, Intense sensation; 6, 7 Very intense sensation. After the bodily sensation scale, participants had the opportunity to describe their bodily sensations in a text box. We calculated a difference score between PAVS and OVS ratings (PAVS–OVS) for pleasantness and bodily sensations to measure the effectiveness of the PAVS stimuli. A larger PAVS–OVS difference score would indicate a greater benefit of PAVS in reducing negative reactions to trigger sounds.

2.4.2. Duke Misophonia Questionnaire (DMQ)

The DMQ is a self-report questionnaire that has been psychometrically validated to measure misophonic experiences across several different dimensions (Rosenthal *et al.*, 2021). The overall DMQ contains eight subscales that measure trigger frequency, affective responses, physiological responses, cognitive responses, coping behaviors before, during, and after being triggered, and beliefs. Participants in our experiment completed all eight subscales; however, we focus our analyses on a composite scale derived from the combined affective, physiological, and cognitive subscales, which we refer to as the *DMQ symptom severity score*. The three subscales asked participants to reflect back on a time where they were intensely bothered by a sound and rate the following: (1) in the affective response subscale (eight items), participants rated how often they felt various emotions such as anger, hostility, and anxiousness; (2) in the physiological response subscale (five items), they rated how often they experienced physiological responses such as becoming rigid, stiff or heart-racing when triggered; (3) in the cognitive subscale (10 items), participants rated how often they had thoughts such as helplessness, wanting to cry, or how to stop the sound. Questions on three subscales were rated on a five-point Likert scale from 0 (Never) to 4 (Always/almost always), and our composite *DMQ symptoms severity score* is computed as the average of all of these items, ranging from 0 to 4.

2.4.3. Amsterdam Misophonia Scale (A-MISO-S)

A-MISO-S is a self-report measure of misophonia that consists of seven items that participants rate on a five-point Likert scale (Schröder *et al.*, 2013). The final item of the questionnaire requires a free text response from the respondent. Scores are categorized into ranges of subclinical (0–4), mild (5–9), moderate (10–14), severe (15–19), and extreme (20–24). This scale measures how much of the participant's time is occupied by trigger sounds, how trigger sounds interfere with social or work life, the amount of distress, the level of distress associated with trigger sounds, the amount of effort to resist thinking about the triggers, and how much control they have over their thoughts about triggers as well as avoidance of trigger sounds.

2.4.4. Misophonia Questionnaire (MQ)

The MQ is a three-part self-report questionnaire that assesses the misophonia symptom's presence, the resulting emotions and behaviors, and the overall severity of sound sensitivity (Wu *et al.*, 2014). The first section, "Misophonia symptoms scale", examines the presence of specific sound sensitivities. Participants rate their sensitivity to certain sounds (e.g., people eating, people making nasal sounds) compared to other people on a scale of 1 (not at all) to 4 (always true). The second section, "Misophonia Emotions and Behaviors", examines emotional and behavioral reactions that arise from being exposed to misophonic trigger sounds. Example items include “Leave the environment to a place where the sound(s) cannot be heard anymore” and “Become anxious or distressed.” Participants rate their experience on a scale of 1 (not at all true) to 4 (always true). The first two sections are combined to a total score ranging from 0 to 68, with higher scores indicating a higher misophonia level. In the last part, “Misophonia severity scale”, participants rated their sound sensitivity on a scale of 0 (minimally) to 15 (very severe). Scores greater than or equal to 7 indicated clinically significant symptoms.

2.4.5. Visual Vividness Imagery Questionnaire (VVIQ)

The VVIQ is a self-report measure of visual imagery, which consists of four sections in which participants are asked to close their eyes and imagine different situations such as “a close friend or relative”, “rising sun”, “a store”, and “a country scene.” There are a total of 16 items that are evaluated on a five-point Likert scale from 1 (No imagery at all, you only “know” that you are thinking of the object) to 5 (Perfectly clear and as lively as normal vision). A total average score is calculated, with higher average scores indicating stronger visual imagery vividness (Marks, 1973).

2.4.6. The Bucknell Auditory Imagery Scale for Vividness (BAIS-V)

The BAIS-V is a self-report measure of auditory imagery vividness (Halpern, 2015). This scale consists of 14 items that describe scenarios involving particular sounds (e.g., happy-birthday song, rainstorm, or dentist’s drill). Participants are asked to read each item and consider whether they think of an auditory image of the described sound in their head. They then rated the vividness of their auditory image on a seven-point Likert scale from 1 (No auditory image is present at all) to 7 (The auditory image is as vivid as the actual sound). Higher scores indicate more vivid auditory imagery.

2.4.7. The Clarity of Auditory Imagery Scale (CAIS)

The CAIS is another self-report scale that measures the clarity of auditory imagery. This scale *consists* of 16 items, such as “a clock ticking”, “a dog barking”, or “paper being torn apart.” Participants are instructed to imagine each item and rate how clearly they could image the sound presented in each item using a five-point Likert scale from 1 (Not clear at all) to 5 (very clear). Higher numbers on this scale indicate more clear auditory imagery (Willander and Baraldi, 2010).

2.5. Qualitative Descriptions of Bodily Sensations

To get a fuller picture of the participant’s experience with the PAVS and OVS stimuli, we examined how each participant described their experience after watching each stimulus. We performed two sets of analyses to analyze these qualitative data. First, we measured the average number of text descriptions as well as the average number of words produced for PAVS and OVS stimuli by misophonia and control participants. Given that misophonia manifests as distinct bodily sensations, we were interested in whether verbal descriptions would include more references to body parts in misophonia participants under the OVS *vs* PAVS context. In order to quantify this, we counted the number of times participants referred to body parts (head, muscle, shoulder, etc.) in their descriptions.

We determined the overall valence of each description by having two independent raters rate each text description based on a three-point valence scale that ranged from positive (+1), to neutral (0), to negative (−1; Cohen's Kappa = 0.79). Text entries that described a pleasant experience, such as feeling relaxed, calm, or happy, were coded as positive, while unpleasant experiences, such as tension, restlessness, or feeling claustrophobic, were coded as negative. Descriptions that could be either positive or negative (e.g., "it was both annoying and amusing") or ambiguous in nature ("tingling sensation") were coded as neutral.

3. Results

3.1. Misophonia Scores

As expected, participants in the misophonia group had significantly higher scores across the three misophonia scales compared to the control group. On the DMQ symptom severity scale, misophonia participants had higher scores ($M = 2.40$, SD = 0.54) compared to the control group ($M = 0.82$, SD = 0.28, $t_{55} = 14.29$, $p < 0.001$, $d = 3.67$, 95% CI [1.37, 1.86]). On the A-MISO-S, misophonia participants also had higher scores ($M = 10.11$, SD = 3.01) compared to the control group ($M = 3.64$, SD = 2.50, $t_{55} = 9.14$, $p < 0.001$, $d = 2.33$, 95% CI [5.05, 7.89]). Misophonia participants also had higher scores on the sensitivity scale of the MQ ($M = 36.73$, SD = 8.29) as well as severity scale ($M = 5.58$, SD = 1.77) compared to the control group ($M = 14.32$, SD = 10.54, $t_{55} = 8.79$, $p < 0.001$, $d = 2.36$, 95% CI [17.29, 27.52]; $M = 2.58$, SD = 2.20, $t_{55} = 5.58$, $p < 0.001$, $d = 1.5$, 95% CI = [1.92, 4.07] respectively).

In this study, the internal consistency (Cronbach's alpha) in the misophonia group was 0.84 for the DMQ symptom severity scale, 0.73 for A-MISO-S, and 0.73 for the sensitivity scale of MQ. In the control group, the Cronbach alpha was 0.91 for the DMQ symptom severity scale, 0.61 for the AMISO-S, and 0.88 for the MQ-Sensitivity scale. Additionally, the three misophonia scales were highly correlated with each other (see Table 5 below). As such, we report results based on DMQ, focusing on the symptom severity subscale, but results are similar across all the scales. Figure 1 shows a histogram distribution of DMQ symptom severity scores for the misophonia and control participants.

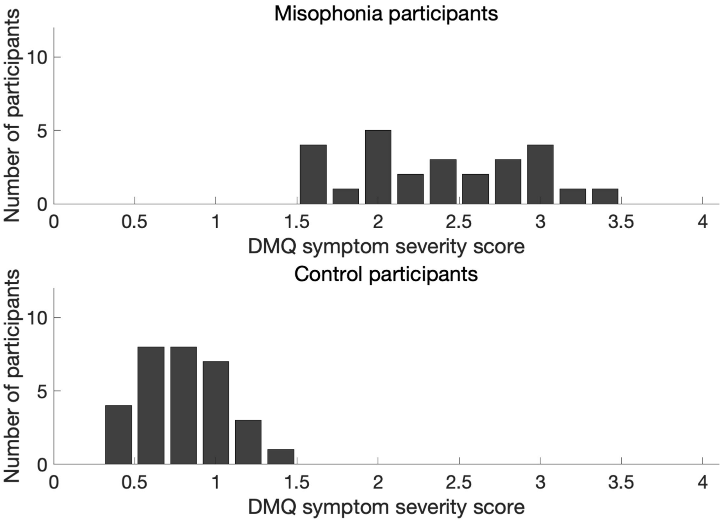

Figure 1. Distribution of Duke Misophonia Questionnaire (DMQ) symptom scale scores for the misophonia and control participants. The misophonia group (top) includes participants ($n = 26$) who scored higher than 1.5 on the DMQ symptom severity scale, and the control group (bottom) includes participants ($n = 31$) who scored lower than 1.5 on the DMQ symptom severity scale.

3.2. Pleasantness Rating for OVS and PAVS Stimuli

To examine how participants in the two groups rated the OVS and PAVS stimuli on the pleasantness scale, we conducted a two-way mixed-effects analysis of variance (ANOVA) with group (misophonia or control) as the between-subjects factor and video type (PAVS or OVS) as the within-subjects factor. There was a significant main effect of group ($F_{1,55} = 20.25$, $p < 0.001$, $\eta_p^2 = 0.32$), reflecting that on average, misophonia participants ($M = 3.07$, SD $=$ 0.50) rated the sounds as significantly less pleasant compared to the control group ($M = 3.62$, SD $= 0.58$). Our analysis also revealed a significant main effect of video type ($F_{1,55} = 162.61$, $p < 0.001$, $\eta_p^2 = 0.56$), reflecting that across all participants, the OVS-paired sounds ($M = 2.87$, SD $= 0.66$) were rated as significantly less pleasant than the PAVS-paired sounds ($M = 3.87$, SD $= 0.56$). There was no statistically significant interaction between group and video type on pleasantness ratings ($F_{1,55} = 0.16$, $p = 0.56$), indicating that the benefit of PAVS over OVS stimuli was similar across misophonia and control participants (see Fig. 2a).

3.3. Bodily Sensation Ratings for OVS and PAVS Stimuli

We also conducted a two-way mixed-effects ANOVA with groups (misophonia or control) as the between-subjects factor and video type (PAVS or OVS)

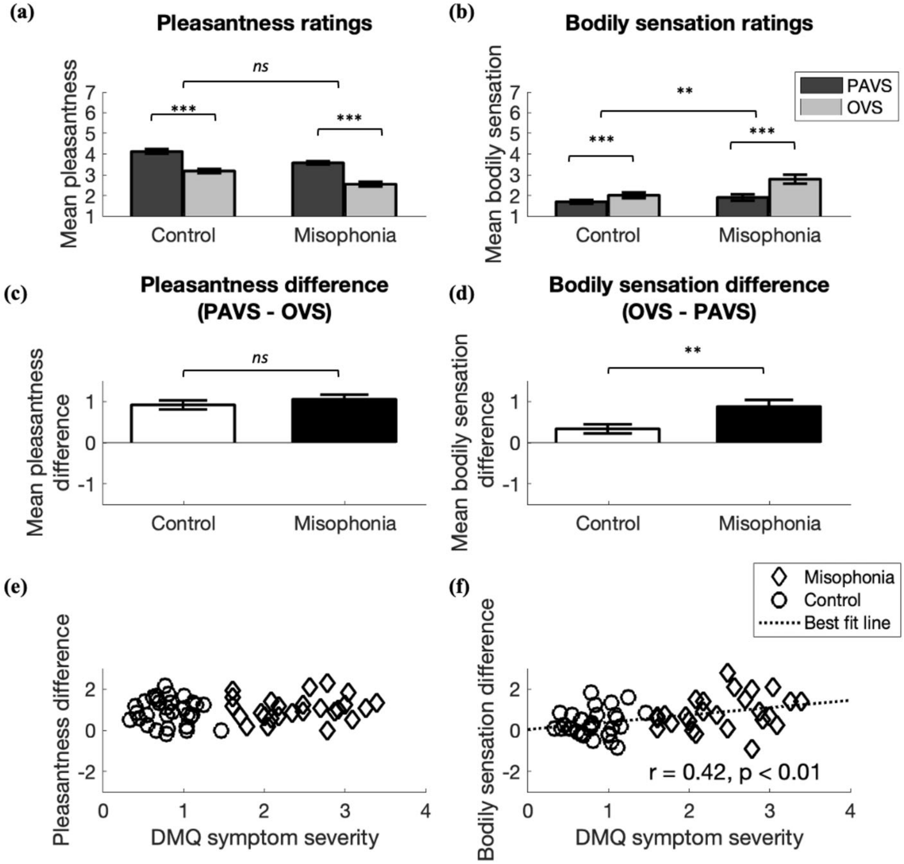

Figure 2. Subjective Reports of Pleasantness and Bodily sensations and Duke Misophonia Questionnaire (DMQ) symptom scale. (a, b) The average ratings for pleasantness and bodily sensations across the control ($n = 31$; light gray bars) and misophonia groups ($n = 26$; dark gray bars). Higher values on the pleasantness scale indicate a more pleasant experience, while a high value on the bodily sensation scale indicates a more intense bodily sensation experience. (c, d) PAVS and OVS difference scores for the control (white bars) and misophonia (black bars) for pleasantness and bodily sensations. (e, f) The correlation between the DMQ symptom severity scale and PAVS–OVS difference scores. Error bars represent the standard error of the mean across participants. ***, $p < 0.001$; **, $p < 0.01$.

as the within-subjects factor to assess the intensity of bodily sensation ratings. There was a significant main effect of group ($F_{1,55} = 6.99$, $p = 0.013$, $\eta_p^2 = 0.11$). On average, the misophonia participants reported experiencing more intense bodily sensations ($M = 2.35$, SD $= 0.96$) compared to the control group ($M = 1.84$, SD $= 0.65$). There was also a significant main effect of video type ($F_{1,55} = 35.14$, p < 0.001, $\eta_p^2 = 0.14$). Across both groups, the OVS stimuli elicited more intense bodily sensations ($M = 2.36$, SD =

1.02) compared to the PAVS stimuli ($M = 1.79$, SD $= 0.65$). Importantly, the interaction between groups and type of video type was statistically significant, $F_{1,55} = 9.59$, $p = 0.005$, $\eta_p^2 = 0.13$, indicating that the benefit of PAVS over OVS in reducing bodily sensations was larger for the misophonia group than for the control group (see Fig. 2b).

To better visualize the benefit of PAVS over OVS stimuli, we calculated a difference score between PAVS and OVS ratings for both the pleasantness and bodily sensation scales for each participant in the misophonia and control groups. A larger PAVS–OVS difference score would indicate a greater benefit from watching PAVS stimuli. Independent-samples t-tests revealed that compared to the control participants, misophonia participants had a significantly larger PAVS–OVS score in bodily sensations rating ($t_{55} = 2.82$, $p = 0.006$, $d = 0.72$, 95% CI [0.15, 0.92]; Fig. 2d), but not in the pleasantness ratings ($t_{55} = 0.68$, $p = 0.39$; Fig. 2c).

We also examined the relationship between misophonia symptoms and the benefits of PAVS over OVS. There was no correlation between DMQ symptom severity scores and PAVS *vs* OVS pleasantness difference scores ($r = 0.12$, $p = 0.37$; Fig. 2e). However, there was a significant positive correlation between DMQ symptom severity scores and bodily sensation difference scores, such that those who scored higher on the DMQ symptom severity scale showed a larger PAVS *vs* OVS difference on the bodily sensation ratings ($r = 0.42$, $p = 0.001$ which survived a Bonferroni correction for two comparisons; Fig. 2f).

3.4. *Item Analysis*

Based on the PAVS and OVS difference scores, we established that 20 out of 26 misophonia participants had at least two PAVS–OVS stimulus pairs that were effective at reducing unpleasantness with a difference score of 2 or above, and 24 out of 26 had two or more stimuli that were effective in reducing bodily sensation to trigger sounds. The specific stimuli that were most effective at reducing bodily sensations and unpleasantness ratings for misophonia participants are shown in Table 3 for misophonia participants and Table 4 for control participants. The full list of video stimuli and their associated difference scores are shown in the Supplementary Fig. S1.

3.5. *Subjective Descriptions of Bodily Sensations*

After rating the intensity of bodily sensations, participants were asked to describe any specific bodily sensations they experienced for each of the 26 stimuli (13 PAVS and 13 OVS). A mixed-effects ANOVA with group (misophonia, control) as the between-subjects factor and stimuli (PAVS, OVS) as the within-subjects factor revealed that overall, misophonia participants produced marginally more descriptions of bodily sensations in response to both

Table 3.
List of stimuli that were most effective for misophonia participants (difference score > 1)

OVS/PAVS stimuli	Pleasantness difference score (PAVS–OVS)	OVS/PAVS stimuli	Bodily sensation difference score (OVS–PAVS)
Gulping/stirring stick in water	2.13	Wet chewing/draining a towel	2.4
Swishing water in mouth/water bottle, shaking	1.4	Clicking pen/bike wheel turning	1.08
Wet chewing/draining a towel	1.25	Fork scraping plate/turning a faucet	1.04
Slurping/water flowing in creek	1.12	Gulping/stirring stick in water	1
Squeaking/puppies barking	1.04	Swishing water in mouth/water bottle shaking	1
Sniffling/pencil drawing on paper	1	–	–
Fork scraping plate/turning a faucet	1	–	–

Table 4.
List of stimuli that were most effective for control participants (difference score > 1)

OVS/PAVS	Pleasantness difference score (PAVS–OVS)	OVS/PAVS	Bodily sensation difference score (OVS–PAVS)
Wet chewing/draining a towel	2.28	Wet chewing/draining a towel	1.68
Gulping/stirring stick in water	1.63	Fork scraping/turning a faucet	1.17
Squeaking/Puppies barking	1.5	–°	–
Fork scraping plate/turning a faucet	1.24	–	–
Sniffling/sweeping floor	1.24	–	–
Sniffling/pencil drawing on paper	1.13	–	–
Swishing water in mouth/water bottle shaking	1.06	–	–

PAVS and OVS stimuli ($M = 6.42$, SD $= 3.54$) compared to the control group ($M = 5.03$, SD $= 3.08$; $F_{1,55} = 3.29$, $p = 0.07$, $\eta_p^2 = 0.06$). There was also a main effect of stimulus type, indicating that both groups of participants produced more descriptions of bodily sensations in response to OVS stimuli ($M = 6.21$, SD $= 3.46$) compared to PAVS stimuli ($M = 5.12$, SD $= 3.19$;

$F_{1,55} = 7.14$, $p = 0.01$, $\eta_p^2 = 0.04$). The interaction between the group and stimulus type was not statistically significant ($F_{1,55} = 2.91$, $p = 0.09$).

To assess the average number of words used by participants per video, a separate mixed-effects ANOVA was conducted with groups as the between-subjects factor and stimulus type as the within-subjects factor. There was no significant main effect of group ($F_{1,55} = 1.29$, $p = 0.26$) or stimuli ($F_{1,55} = 0.80$, $p = 0.37$); however, there was a significant interaction ($F_{1,55} = 3.22$, $p = 0.07$, $\eta_p^2 = 0.05$), suggesting that the difference between words used per video for OVS and PAVS stimuli was larger in the misophonia group. On average, misophonia participants used more words per video to describe their experience with OVS ($M = 6.42$, SD $= 5.62$) than PAVS stimuli ($M = 4.95$, SD $= 5.74$).

Moreover, we tested how often descriptions referred to specific body parts with a mixed-effects ANOVA with groups (misophonia, control) as the between-subjects factor and stimulus type as the within-subjects variable. There was a significant main effect of groups, indicating that, on average, misophonia participants made more references to specific body parts when describing their experience with the PAVS and OVS stimuli ($M = 8.02$, SD $=$ 7.54) compared to the control group ($M = 3.31$, SD $= 4.39$; $F_{1,55} = 12.67$, $p < 0.001$, $\eta_p^2 = 0.19$). There was also a main effect of stimulus type, suggesting that all participants made more references to body parts when describing their experience with OVS stimuli ($M = 7.53$, SD $= 7.11$) compared to PAVS stimuli ($M = 3.39$, SD $= 4.98$; $F_{1,55} = 31.02$, $p < 0.001$, $\eta_p^2 = 0.15$). The interaction between the group and stimuli was not statistically significant, $F_{1,55} = 0.025$, $p = 0.87$. When controlling for the number of words each participant used, the mixed-effects ANOVA revealed that misophonia participants referenced body parts more frequently compared to the control participants ($F_{1,55} = 10.11$, $p = 0.002$, $\eta_p^2 = 0.15$). Additionally, all participants reported body parts more frequently to OVS stimuli ($M = 0.14$, SD $= 0.12$) compared to PAVS stimuli ($M = 0.07$, SD $= 0.09$; $F_{1,55} = 16.45$, $p < 0.001$, $\eta_p^2 = 0.15$). A significant interaction also revealed that the difference between body part frequency between OVS stimuli ($M = 0.11$, SD $= 0.10$) and PAVS stimuli ($M = 0.01$, SD $= 0.03$) was larger for the control group compared to the misophonia group ($F_{1,55} = 4.74$, $p = 0.03$, $\eta_p^2 = 0.08$).

To assess the valence of descriptions, a mixed-effects ANOVA was conducted with group (misophonia and control) as the between-subjects factor, stimulus type (PAVS or OVS) as the within-subjects factor, and the judged valence of the descriptions as the dependent measure (see Section 2. *Method*). There was a main effect of the participant group, $F_{1,24} = 12.04$, $p < 0.01$, $\eta_p^2 = 0.33$, indicating that misophonia participants used more negative descriptions overall ($M = -0.64$, SD $= 0.25$) compared to the control group ($M = -0.39$, SD $= 0.33$). There was also a main effect of stimulus type

($F_{1,24} = 18.74$, $p < 0.001$, $\eta_p^2 = 0.42$), which shows that overall OVS descriptions were rated as more negative ($M = -0.67$, $SD = 0.27$) compared to PAVS descriptions ($M = -0.37$, $SD = 0.28$). The interaction between the participant's group and stimulus type was not statistically significant ($F_{1,24} = 0.016$, $p = 0.89$). Example quotes from participants with misphonia powerfully illustrate this effect. For instance, one misophonia participant described their bodily sensation to the OVS video of gulping as "my nose scrunched up, and I felt some tension in my body,— more from the 'gulping' sounds than the sound of the water flowing through the water bottle." The same participant described their bodily sensation to the same sound paired with the PAVS video (gulping sound paired with video of flowing water in a creek) as "relaxation in my chest." Similarly, another misophonia participant described their bodily sensation to the OVS-paired fork tapping sound as "felt sharp in my chest." but described the bodily sensation to the same sound paired with the PAVS video (fork tapping sound paired with hammering nail video) as "felt it in the base of my skull and in my cheeks. The sound felt light and happy."

3.6. Visual and Auditory Imagery

An independent-samples t-test showed that the clarity of auditory imagery (CAIS) scores for the misophonia group ($M = 4.11$, $SD = 0.74$) was significantly higher than for the control group ($M = 3.62$, $SD = 0.75$; $t_{55} = 2.45$, $p = 017$, $d = 0.66$, 95% CI [0.09, 0.89]). However, BAIS-V scores for the misophonia group ($M = 4.65$, $SD = 1.36$) were not significantly different from those for the control group ($M = 4.11$, $SD = 1.34$; $t_{55} = 1.50$, $p = 0.14$). Additionally, VVIQ scores for the misophonia group ($M = 3.27$, $SD = 0.58$) were not significantly different from those for the control group ($M = 3.12$, $SD = 0.70$; $t_{55} = 0.91$, $p = 0.37$).

Furthermore, CAIS scores were significantly correlated with DMQ symptom severity scores ($r = 0.42$, $p < 0.001$), indicating that those who scored higher on the DMQ also reported having more clear auditory imagery. The vividness of auditory imagery (BAIS-V) was marginally correlated with DMQ symptom severity scores ($r = 0.25$, $p = 0.052$). However, visual imagery (VVIQ) was not correlated with DMQ symptom severity scores ($r = 0.046$, $p = 0.73$). Neither the clarity of auditory imagery (CAIS), the vividness of auditory imagery (BAIS-V), nor the VVIQ scores were correlated with PAVS–OVS difference scores on bodily sensation scales or pleasantness ratings (all r values < 0.24, and all p values > 0.1). Overall, participants with higher DMQ symptom scores tended to have clearer auditory imagery, yet the clarity of auditory imagery was not predictive of how well the PAVS stimuli reduced bodily sensations compared to the OVS stimuli. We note that these comparisons were not corrected for multiple comparisons. A Bonferroni correction

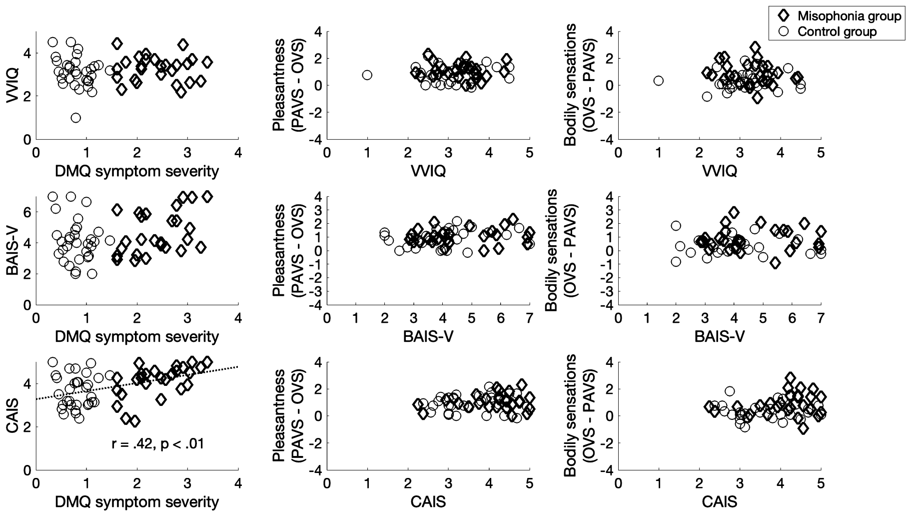

Figure 3. Correlation between visual, auditory, and Duke Misophonia Questionnaire (DMQ) symptom severity scales. Pearson correlations between DMQ symptom scores and Vividness of Visual Imagery Questionnaire (VVIQ), Bucknell Auditory Imagery Scale for Vividness (BAIS-V), and the Clarity of Auditory Imagery Scale (CAIS), as well as PAVS–OVS ratings across the pleasantness and bodily sensation scales. Best-fit lines were added for a significant correlation between variables.

considering the nine different comparisons made above would imply a threshold of alpha = 0.005. The correlation between CAIS and DMQ scores survived this correction. Figure 3 shows graphs for the above correlations, and Table 5 shows the associated *r* and *p* values.

4. Discussion

In this study, we solidified the notion that reactions to misophonic trigger sounds are malleable and susceptible to higher-level cognitive influences. Importantly, our results demonstrated the multisensory nature of the misophonic experience. We showed that PAVS synchronized with common trigger sounds can significantly reduce feelings of unpleasantness and the intensity of bodily sensations in both a misophonia and a control sample. This finding extends our previous results on a neurotypical population (Samermit *et al.*, 2022), showing that sound-swapped video stimuli are just as effective at reducing negative reactions in people with misophonia. In fact, the current study found that people in the misophonia group experienced more relief in bodily sensations from the PAVS stimuli than control participants, whereas

Table 5.
Correlation matrix between all ratings

	DMQ symptom severity	MQ sensitivity	MQ severity	A-MISO-S	Pleasantness difference	Bodily sensation difference	VVIQ	BAIS-V	CAIS
DMQ symptom severity	–	–	–	–	–	–	–	–	–
MQ sensitivity	**0.85*** **	–	–	–	–	–	–	–	–
MQ severity	**0.60*** **	**0.70*** **	–	–	–	–	–	–	–
A-MISO-S	**0.78*** **	**0.77*** **	**0.78*** **	–	–	–	–	–	–
Pleasantness difference	0.12	0.06	0.12	0.23^{+}	–	–	–	–	–
Bodily sensation difference	**0.42** **	**0.38** **	**0.31***	**0.47*** **	**0.30***	–	–	–	–
VVIQ	0.04	−0.01	−0.07	0.01	0.06	−0.07	–	–	–
BAIS-V	0.25^{+}	−0.15	0.02	0.09	0.22	0.08	**0.62** **	–	–
CAIS	**0.43** **	**0.35** **	0.12	0.23	0.13	0.24	**0.42** **	**0.67*** **	–

Notes. Significant correlations are bolded. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$, $^{+}$, marginally significant.

both groups experienced similar effects on pleasantness ratings. This interaction indicates that reported bodily sensations may be more representative of the misophonic experience than pleasantness ratings.

The majority of the misophonia participants had at least two sound-swapped video pairs that were effective at reducing misophonic reactions, which indicates that our stimuli set is adequate to cover most cases of misophonia. Our rating analyses are supported by open-ended text descriptions provided by participants about their qualitative experience participants in response to the stimuli. Misophonia participants not only used more words to describe their bodily sensations in response to each stimulus, but they also referenced specific body parts more frequently. In addition, their descriptions of bodily sensations were more negative in nature compared to the control group.

Additionally, we examined whether other traits, such as visual or auditory imagery, differed among people with and without misophonia and whether they predict the benefits of PAVS stimuli. We reasoned that people with more vivid auditory and visual imagery might be able to better ‘perceptually map’ the OVS sound with the PAVS video. Previous research has shown people with misophonia report having more clear and vivid auditory imagery compared to people without misophonia (Simner *et al.*, 2021). Our results confirm this prior finding, with participants in the misophonia group experiencing clearer auditory imagery than those in the control group. Perhaps heightened auditory imagery is associated with heightened auditory attention, thereby making the trigger sounds more salient for people with misophonia. However, in our sample, visual imagery abilities were not different between misophonia and control groups.

Additionally, we can speculate that the effectiveness of the PAVS stimuli when the true source is known may be related to the ability to imagine alternative sources. That is, when someone actually knows that they are hearing a crunching chewing sound but are watching a piece of paper being torn, in order for it to be effective at reducing the negative reaction to the chewing sound, they need to suppress the representation of the true source (i.e., someone chewing) and allow the PAVS (i.e., piece of paper being torn) to override the knowledge of the true source. This effect might be linked to the ability to conjure visual and auditory imagery. In our experiment, we found that misophonia participants have at least as good visual imagery as controls and have more clear auditory imagery ability than controls according to one scale (CAIS) but not another (BAIS). Although there was no relationship between imagery skills and the PAVS benefits, we believe this is a promising direction for potential interventions, as people with misophonia can potentially leverage their imagery abilities to imagine alternative sources for a sound when they are triggered, and a PAVS is not physically available. For example, if a participant finds it effective as a way to attenuate their reaction to chewing, perhaps the simple act of imagining someone walking on leaves can have similar effects.

While the current study used multisensory stimuli to modify the belief about the source of a trigger sound, we cannot be sure that other ways of manipulating beliefs about the sound source, such as expectations (Heller and Smith, 2022) or reduced certainly about the source (Savard *et al.*, 2022) cannot have a similar effect in reducing negative reactions to trigger sound. In order to conclude that our manipulation is truly multisensory, a future study would need to test whether the effects depend on the audiovisual synchrony of the stimuli. For audiovisual integration to happen, auditory and visual stimuli need to occur within a certain period of time, a window of time referred to as the temporal binding window (TBW) of sensory integration (Stevenson *et al.*, 2014). If audiovisual stimuli are desynchronized enough to fall outside of the TBW, sensory integration is interrupted, and the audiovisual stimuli are perceived as two separate events. As such, if the PAVS–OVS benefit does not manifest when the audio and video streams are desynchronized, then we would be able to conclude that multisensory integration is fundamental. Alternatively, if effects are similar regardless of audiovisual synchrony, or if they manifest similarly with a video, a still picture, or verbal descriptions of the alternative source, then we would conclude that the effects stem from higher-level cognitive mechanisms and multisensory stimuli just happen to be an effective way to engage the system.

Although we do not have direct evidence, our interview with misophonic participants (Samermit *et al.*, 2022) revealed that a substantial proportion of misophonia participants reported worse reactions when they could see the source of the sound compared to when they could not. In these real-life contexts, seeing *vs* not seeing the source is unlikely to affect beliefs about the nature of the sound source; instead, it seems more likely that seeing the source makes the response worse because it enhances the multisensory processing of the trigger action (e.g., chewing, gulping, sniffling, etc.).

An important limitation of the current study is that we did not account for other disorders that are comorbid with misophonia. A number of studies have shown that misophonia generally co-occurs with anxiety, OCD, PTSD, depression, and autistic traits (Edelstein *et al.*, 2013; Erturk *et al.*, 2023; Jager *et al.*, 2020; Rinaldi *et al.*, 2023; Rouw and Erfanian, 2018; though Schröder *et al.*, 2013; see Smit *et al.*, 2023 for contrasting results regarding autism). As such, we cannot be sure that the differences between groups are not driven by other attributes that are known to be comorbid with misophonia. For example, the greater reduction in bodily sensation experienced by misophonia participants compared to controls may have been driven by differences in anxiety or depression levels.

Additionally, in our current study, we did not manipulate the presentation order of PAVS and OVS stimuli. Participants always watched and rated the 13 PAVS videos (in a random order) and then watched the 13 corresponding OVS videos in a new random order. Previously, Samermit and colleagues (2022) showed that watching the OVS stimuli first attenuates the benefits of PAVS stimuli. In other words, exposure to the true source of a trigger sound made a subsequent positive alternative visual source less effective in reducing negative reactions. In the Samermit *et al.* (2022) study, the PAVS–OVS benefit when the PAVS stimuli were presented first was higher (0.65 difference within a five-point scale) than when the OVS stimuli were presented first (0.39 difference). That is, when the OVS stimuli were presented first, the PAVS–OVS benefit was approximately 60% as strong as when the PAVS stimuli were presented first. As such, while the fixed order in the current study might contribute to an overestimation of the PAVS *vs* OVS effect, the effect is not completely eliminated when the OVS stimuli are presented first. It is important to note that we chose this fixed order as a stepping stone for a future study that aims to investigate the effectiveness of these stimuli as an intervention in a longitudinal study. This allowed us to identify stimuli that are most effective at reducing negative reactions to trigger sounds when the participant is unaware of the true source. Future research still needs to address the degree to which knowledge of the true trigger source dampens the effectiveness of the PAVS stimuli.

Given that subjective reports of bodily sensations were a representation of the benefit of watching PAVS stimuli, distinguishing participants with misophonia from controls, future studies should incorporate these and additional measures of bodily responses, such as GSR, electromyography (EMG), as well as pupillary responses. While some recent studies utilizing such psychophysiological methods have shown increased sympathetic activation in response to a trigger and aversive sounds in misophonic participants (Grossini *et al.*, 2022), others have found similar psychophysiological responses to trigger sounds in both misophonia and control participants despite reporting subjective differences (Neacsiu *et al.*, 2023). It would be beneficial to understand if the subjective ratings of reduced bodily sensations to PAVS stimuli, which were more pronounced for our misophonia participants, would manifest in psychophysiological measures. Research on the multisensory and top-down influences in misophonia may pave the way for new interventions for misophonia sufferers.

Implications of these findings include potential interventions for misophonia sufferers involving a type of exposure therapy (ET) to mitigate reactions to trigger sounds. While prior studies have shown that ET can be effective for a number of mental health disorders, including OCD (Koran *et al.*, 2007), PTSD (Bradley *et al.*, 2005), and anxiety (Rodebaugh *et al.*, 2004), repeated exposure to trigger sounds has been ineffective in misophonia, likely due to the intensity of reactions to trigger stimuli, which discourages misophonics to continue with treatment (Frank and McKay, 2019; Schröder *et al.*, 2017). Our PAVS stimuli provide a way to expose misophonic participants to trigger sounds in a more tolerable way. Another potential implication of our finding involves interventions using mental imagery of an alternative positive visual source in the presence of a trigger sound. This could be especially helpful in a real-life context where a trigger sound is present, but the usual coping mechanisms (e.g., leaving, putting on headphones, etc.) or videos of alternative visual sources are not appropriate or available. Conjuring up mental imagery of a positive alternative source could potentially provide a similar benefit to what was observed in our lab study in real-life situations. While such an intervention is promising, it should be noted that not everyone has similar imagery abilities; as such, it is possible that this might only be useful for misophonia sufferers who have higher imagery abilities.

Supplementary Material

Supplementary material is available online at:
https://doi.org/10.6084/m9.figshare.27609624

References

Berger, C. C. and Ehrsson, H. H. (2013). Mental imagery changes multisensory perception, *Curr. Biol.* **23**, 1367–1372. DOI:10.1016/j.cub.2013.06.012.

Berger, C. C. and Ehrsson, H. H. (2014). The fusion of mental imagery and sensation in the temporal association cortex. *J. Neurosci.* **34**, 13684–13692.

Berger, C. C. and Ehrsson, H. H. (2018). Mental imagery induces cross-modal sensory plasticity and changes future auditory perception. *Psychol. Sci.* **29**, 926–935.

Bradley, R., Greene, J., Russ, E., Dutra, L. and Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD, *Am. J. Psychiatry* **162**, 214–227. DOI:10.1176/appi.ajp.162.2.214.

Dixon, L. J., Schadegg, M. J., Clark, H. L., Sevier, C. J. and Witcraft, S. M. (2024). Prevalence, phenomenology, and impact of misophonia in a nationally representative sample of US adults. *J. Psychopathol. Clin. Sci.* **133**, 403–412. DOI:10.1037/abn0000904.

Edelstein, M., Brang, D., Rouw, R. and Ramachandran, V. S. (2013). Misophonia: physiological investigations and case descriptions, *Front. Hum. Neurosci.* **7**, 296. DOI:10.3389/fnhum.2013.00296.

Ertuerk, E., Işik, Ü., Aktepe, E. and Kiliç, F. (2023). Examining the correlation between misophonia symptoms and autistic traits in general population. *Int. J. Dev. Disabil.* DOI:10.1080/20473869.2023.2165231.

Frank, B. and McKay, D. (2019). The suitability of an inhibitory learning approach in exposure when habituation fails: a clinical application to misophonia, *Cogn. Behav. Pract.* **26**, 130–142. DOI:10.1016/j.cbpra.2018.04.003.

Grossini, E., Stecco, A., Gramaglia, C., De Zanet, D., Cantello, R., Gori, B., Negroni, D., Azzolina, D., Ferrante, D., Feggi, A., Carriero, A. and Zeppegno, P. (2022). Misophonia: analysis of the neuroanatomic patterns at the basis of psychiatric symptoms and changes of the orthosympathetic/parasympathetic balance, *Front. Neurosci.* **16**, 827998. DOI:10.3389/fnins.2022.827998.

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes, *Psychomusicology* **25**, 37–47. DOI:10.1037/pmu0000081.

Heller, L. M. and Smith, J. M. (2022). Identification of everyday sounds affects their pleasantness, *Front. Psychol.* **13**, 894034. DOI:10.3389/fpsyg.2022.894034.

Jager, I., de Koning, P., Bost, T., Denys, D. and Vulink, N. (2020). Misophonia: phenomenology, comorbidity and demographics in a large sample, *PLoS One* **15**, e0231390. DOI:10.1371/journal.pone.0231390.

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

Jastreboff, P. J. and Jastreboff, M. M. (2014). Treatments for decreased sound tolerance (hyperacusis and misophonia), *Semin. Hear.* **35**, 105–120. DOI:10.1055/s-0034-1372527.

Koran, L. M., Hanna, G. L., Hollander, E., Nestadt, G., Simpson, H. B. and American Psychiatric Association (2007). Practice guideline for the treatment of patients with obsessive-compulsive disorder. *Am. J. Psychiatry* **164**, 5–53.

Kosslyn, S. M., Thompson, W. L., Klm, I. J. and Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature* **378**, 496–498.

Kumar, S., Tansley-Hancock, O., Sedley, W., Winston, J. S., Callaghan, M. F., Allen, M., Cope, T. E., Gander, P. E., Bamiou, D.-E. and Griffiths, T. D. (2017). The brain basis for misophonia, *Curr. Biol.* **27**, 527–533. DOI:10.1016/j.cub.2016.12.048.

Pearson, J., Clifford, C. W. and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* **18**, 982–986.

Marks, D. F. (1973). Vividness of Visual Imagery Questionnaire, *J. Ment. Imag.* DOI:10.1037/t05959-000.

Neacsiu, A. D., Beynel, L., Gerlus, N., Szymkiewicz, V., LaBar, K. S., Bukhari-Parlakturk, N. and Rosenthal, M. Z. (2023). Identifying optimal parameters for neuroscience-informed interventions for misophonia, medRxiv, 2023-06. DOI:10.1101/2023.06.25.23291872.

Rinaldi, L. J., Simner, J., Koursarou, S. and Ward, J. (2023). Autistic traits, emotion regulation, and sensory sensitivities in children and adults with Misophonia, *J. Autism Dev. Disord.* **53**, 1162–1174. DOI:10.1007/s10803-022-05623-x.

Rodebaugh, T. L., Holaway, R. M. and Heimberg, R. G. (2004). The treatment of social anxiety disorder, *Clin. Psychol. Rev.* **24**, 883–908. DOI:10.1016/j.cpr.2004.07.007.

Rosenthal, M. Z., Anand, D., Cassiello-Robbins, C., Williams, Z. J., Guetta, R. E., Trumbull, J. and Kelley, L. D. (2021). Development and initial validation of the Duke Misophonia Questionnaire, *Front. Psychol.* **12**, 709928. DOI:10.3389/fpsyg.2021.709928.

Rouw, R. and Erfanian, M. (2018). A large-scale study of misophonia, *J. Clin. Psychol.* **74**, 453–479. DOI:10.1002/jclp.22500.

Samermit, P., Saal, J. and Davidenko, N. (2019). Cross-sensory stimuli modulate reactions to aversive sounds, *Multisens. Res.* **32**, 197–213. DOI:10.1163/22134808-20191344.

Samermit, P., Young, M., Allen, A. K., Trillo, H., Shankar, S., Klein, A., Kay, C., Mahzouni, G., Reddy, V., Hamilton, V. and Davidenko, N. (2022). Development and evaluation of a sound-swapped video database for misophonia, *Front. Psychol.* **13**, 890829. DOI:10.3389/fpsyg.2022.890829.

Savard, M. A., Sares, A. G., Coffey, E. B. and Deroche, M. L. D. (2022). Specificity of affective responses in misophonia depends on trigger identification, *Front. Neurosci.* **16**, 879583. DOI:10.3389/fnins.2022.879583.

Schröder, A., Vulink, N. and Denys, D. (2013). Misophonia: diagnostic criteria for a new psychiatric disorder, *PLoS One* **8**, e54706. DOI:10.1371/journal.pone.0054706.

Schröder, A. E., Vulink, N. C., van Loon, A. J. and Denys, D. A. (2017). Cognitive behavioral therapy is effective in misophonia: an open trial, *J. Affect. Dis.* **217**, 289–294. DOI:10.1016/j.jad.2017.04.017.

Siepsiak, M., Vrana, S. R., Rynkiewicz, A., Rosenthal, M. Z. and Dragan, W. Ł. (2022). Does context matter in misophonia? A multi-method experimental investigation, *Front. Neurosci.* **16**, 880853. DOI:10.3389/fnins.2022.880853.

Simner, J., Koursarou, S., Rinaldi, L. J. and Ward, J. (2021). Attention, flexibility, and imagery in misophonia: does attention exacerbate everyday disliking of sound?, *J. Clin. Exp. Neuropsychol.* **43**, 1006–1017. DOI:10.1080/13803395.2022.2056581.

Smit, D. J. A., Bakker, M., Abdellaoui, A., Hoetink, A. E., Vulink, N. and Denys, D. (2023). A genome-wide association study of a rage-related misophonia symptom and the genetic link with audiological traits, psychiatric disorders, and personality, *Front. Neurosci.* **16**, 971752. DOI:10.3389/fnins.2022.971752.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M. and Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders, *J. Neurosci.* **34**, 691–697. DOI:10.1523/JNEUROSCI.3615-13.2014.

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. and Raver, S. M. (2022). Consensus definition of misophonia: a Delphi study, *Front. Neurosci.* **16**, 841816. DOI:10.3389/fnins.2022.841816.

Vitoratou, S., Uglik-Marucha, N., Hayes, C. and Gregory, J. (2021). Listening to people with misophonia: exploring the multiple dimensions of sound intolerance using a new psychometric tool, the S-five, in a large sample of individuals identifying with the condition, *Psych* **3**, 639–662. DOI:10.3390/psych3040041.

Willander, J. and Baraldi, S. (2010). Development of a new clarity of auditory imagery scale, *Behav. Res. Methods* **42**, 785–790. DOI:10.3758/BRM.42.3.785.

Wu, M. S., Lewin, A. B., Murphy, T. K. and Storch, E. A. (2014). Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. *J. Clin. Psychol.* **70**, 994–1007.