

Examining Physiological Responses to Misophonic Triggers

C. O'Reilly^{1,4}, X. Yang^{2,5}, S. Oh^{2,5}, D. Wedell^{2,5}, and S. V. Shinkareva^{2,5}

¹ Department of Computer Science and Engineering, University of South Carolina, Columbia, SC, USA

² Institute for Mind and Brain, University of South Carolina, Columbia, SC, USA

³ Artificial Intelligence Institute, University of South Carolina, Columbia, SC, USA

⁴ Carolina Autism and Neurodevelopment Research Center, University of South Carolina, Columbia, SC, USA

⁵ Department of Psychology, University of South Carolina, Columbia, SC, USA

E-mail: christian.oreilly@sc.edu; shinkareva@sc.edu

Summary: We collected and analyzed face electromyogram, skin electrodermal activity, peripheral skin temperature, and electrocardiogram in 60 participants with (N=35) and without (N=25) misophonia, a condition characterized by decreased tolerance to innocuous sounds. Our goal was to characterize the physiological response to misophonia-triggering sounds objectively. We found that misophonic responses can be identified in some cases through atypical physiological reactions to triggering stimuli, though not all participants exhibited this response. Our analyses suggest a large interindividual variability in response to misophonic triggers and highlight the need for methodological adjustments in future experiments to increase the detectability of misophonic reactions.

Keywords: biosignals, misophonia, electromyogram, electrodermal activity, electrocardiogram, peripheral skin temperature, machine learning.

1. Introduction

Misophonia is a condition characterized by a decreased tolerance to innocuous sounds (i.e., triggers), such as chewing, which can be highly debilitating for some individuals [1]. Nearly 5% of the adult population is affected by misophonia [2, 3]. Despite its prevalence, research on misophonia remains limited, and the biological mechanisms underlying this condition are poorly understood. This gap in knowledge hinders the development of effective treatments and preventive interventions. Current approaches to studying misophonia often rely on participants' self-report ratings of distress in response to triggers, which are subjective and prone to variability within participants. To address this issue, identifying misophonic triggers through physiological signals offers a more objective means of confirming responses and assessing the severity of this condition. Electrodermal activity (EDA) and heart rate have previously been used to describe misophonic responses [4-7]. This study aims to further characterize the physiological responses associated with misophonia by also including facial electromyography (EMG) and peripheral skin temperature (SKT), two recording modalities that have not been previously explored in this condition.

2. Methods

Experimental paradigm. We recorded five physiological signals: EMG for the corrugator supercilii (EMG C) and zygomaticus major (EMG Z) muscles, EDA, SKT (measured on the palm side of the thumb of the left hand), and electrocardiogram (ECG), at 1,000 Hz¹ [8] in participants with and without misophonia. Participants listened to sounds, viewed silent videos associated with sounds, or were asked to think about sounds in a 3 (sensory modality: auditory, visual, mental imagery) by 3 (stimuli:

trigger, aversive, non-aversive) factorial design. Participants rated how distressing (misophonia group) or anti-social (control group) the stimulus was for each trial (Fig. 1). This distinction between groups is necessary because distress is a key feature of misophonic responses, which is not typically experienced by control participants, even though they may find the stimuli unpleasant or antisocial. Participants aged 18 to 45 without a history of neurological or psychiatric diagnoses, hearing loss, or hyperacusis were drawn from the South Carolina community and compensated for their time. The Misophonia group (N=35, 30 female; M_{age}=25.3, SD_{age}=8.10) consisted of individuals who experience misophonic reactions triggered by sounds with distinct visual components but not associated with specific individuals. Trigger sets for each misophonia participant were selected based on the categories of triggers identified during the screening interviews. Control group participants (N=25, 20 female; M_{age}=25.0, SD_{age}=8.10; the recruitment of control participants is ongoing) were matched with misophonia participants on age, biological sex, handedness, and stimuli presentation. Control participants were also screened to ensure they did not have misophonia. All participants completed two misophonia severity questionnaires, the Duke-Vanderbilt Misophonia Screening Questionnaire (DVMSQ) [9] and the Selective Sound Sensitivity Syndrome Scale (S-Five) [3]. They also filled two mental imagery questionnaires, the Vividness of Visual Imagery Questionnaire (VVIQ) [10] and the Bucknell Auditory Imagery Scale (BAIS) [11], which included two subscales for vividness (BAIS-V) and control (BAIS-C). The study was carried out in accordance with the procedures and protocols approved by the University of South Carolina Institutional Review Board, and all participants signed an informed consent.

¹ Before preprocessing, six recordings acquired at 200 Hz and two at 2,000 Hz were resampled at 1,000 Hz using the MNE-Python `resample()` function, which

adopts the same approach as SciPy, relying on the Fast Fourier Transform. We observed no systematic effects of this resampling on classification.

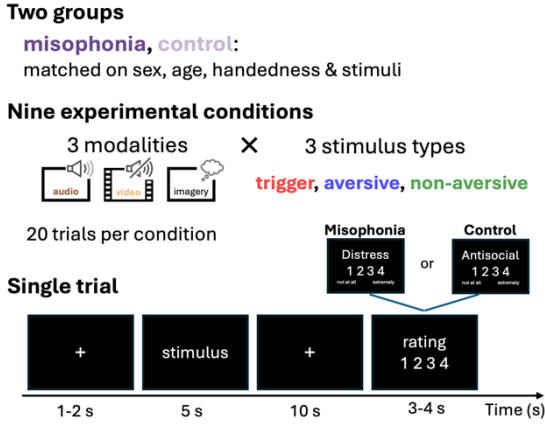


Fig. 1. Experimental design.

Preprocessing and feature computation. EMG signals were notch-filtered at 60 Hz and band-pass filtered in the 20-500 Hz range. They were then rectified, baseline-corrected using the average amplitude over one second before the stimulus onset, and smoothed with a rolling average (window size = 100 samples, or 0.1 s). From these preprocessed signals, we extracted the peak amplitude within the 7-second window following the stimulus as features for classification. EDA and temperature were similarly baseline-corrected and smoothed. For these signals, the amplitude 15 seconds after stimulus onset was used as a feature. For the ECG signal, R peaks were automatically detected using AcqKnowledge (Biopac Systems, Inc., Goleta, CA, USA). Two research assistants manually edited the R peaks to correct motion artifacts and misclassifications. The instantaneous heart rate was computed as the inverse of the R-R interval and used to characterize the ECG signals. We used the mean-square-root of the baseline-corrected heart rate within a 2-7 second window following stimulus onset as a feature. The defined windows were based on preliminary analyses. EMG Z data from one control participant were excluded due to issues related to the data acquisition process.

Classical statistical analysis. Student's t-tests and correlation analysis were used to compare stimulus types within individuals for each physiological signal. The correlation between ratings and physiological responses was assessed using Pearson's correlation across all trials, regardless of stimulus type or sensory modality. All tests were conducted at an alpha level of 0.001 without correction for multiple comparisons. For statistical significance between time series, multiple comparisons were corrected using a cluster-level statistical permutation test, as implemented in MNE-Python [12]. Since not all trigger trials led to misophonic responses, we selected only half of the trigger trials based on the higher distress/antisocial ratings reported by the participants. This selection ensures that the average physiological response to triggers is not diluted by trials in which participants did not subjectively feel triggered. Since preliminary analyses showed a weaker response for the mental imagery, it has been excluded from the classical statistical analyses. We report on the relative strength of responses by modality in the machine learning analysis.

Machine learning analysis. We used machine learning to determine whether physiological responses could

predict group membership (i.e., misophonia versus control) and stimulus type (i.e., trigger, aversive, and non-aversive). Linear Support Vector Machines with default parameterization were employed (using the LinearSVC implementation in Scikit-Learn [13], based on LIBLINEAR [14]). A leave-one-out cross-validation was used to assess group classification, and a leave-one-group-out cross-validation, with the participant serving as the grouping factor, was used for assessing stimulus type classification. Features were standardized by removing the mean and scaling to unit variance before classification. For stimulus classification, both accuracy and weighted f1 scores were used to measure performance for the balanced case. For the unbalanced case of group classification, only weighted f1 scores were reported. To assess the statistical significance of f1 scores and accuracies, we bootstrapped the trial selection within participants 100 times using random selection with full sample size and replacement. We evaluated the relative importance of the different features in the classification using permutation feature importance [15]. We imputed missing data (i.e., features from the excluded EMG Z channel for one control participant) using the nearest neighbors imputation approach [16] provided by the scikit-learn KNNImputer class.

3. Results

First, we examined individual differences in physiological responses to trigger versus aversive stimuli for each signal. We found that misophonia triggers could be identified in some participants (Fig. 2A) and at the group level (Fig. 2D) based on physiological responses. Participants who exhibited significant physiological responses to triggers were generally those who self-reported experiencing the most distress when presented with these triggers. This is demonstrated by the fact that all but one significant physiological response differences at the individual level (Fig. 2A) were observed in the nine participants who showed a clear difference in ratings between aversive and trigger stimuli, with t-statistics greater than 10 (Fig. 2C). Additionally, self-report and physiological responses were strongly correlated in a large proportion of participants, particularly within the misophonia group (Fig. 2B).

To further investigate the predictive value of physiological responses, we used machine learning to assess whether group membership and stimulus types could be predicted from physiological data. When all features of the physiological responses were considered together, participant classification into misophonia and control groups was not statistically significant. Not surprisingly, self-report distress/antisocial ratings were strong predictors of group membership (Fig. 3A). However, predictive information was found in physiological responses when contrasts between stimulus types were used (Fig. 3B). The contrast between aversive and non-aversive stimuli appeared to be as indicative of misophonia as contrasts involving trigger stimuli. Additionally, mental imagery contrasts between stimulus types seemed to provide as much predictive information as perceived auditory or visual stimuli.

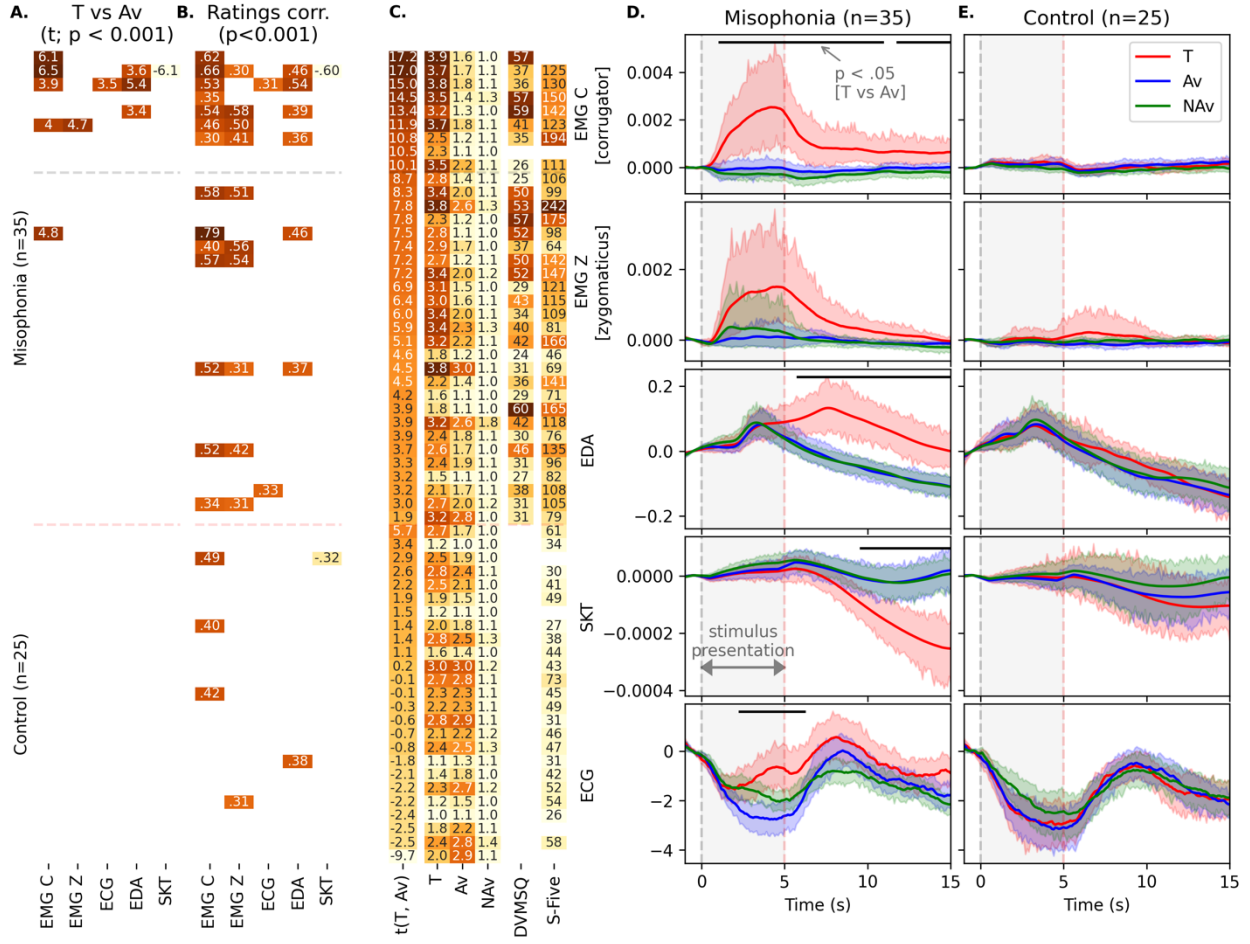


Fig. 2. Detecting triggers from physiology. **A)** Significant t values from Student t -tests comparing trigger and aversive responses by physiological signal and participant. **B)** Significant Pearson's correlations between physiological responses and distress (misophonia) or antisocial (control) ratings. **C)** Column 1: Student's t values for the difference in means between distress/antisocial ratings for triggers (T) and aversive (Av) stimuli, noted $t(T, Av)$. Columns 2-4: Average distress/antisocial ratings for T, Av, and NAv (non-aversive) stimuli. Columns 5-6: DVMSQ and S-Five misophonia severity scores. For A-C, the dashed light red line separates misophonia and control participants. Within groups, participants are sorted by decreasing $t(T, Av)$ values. The dashed light gray line identifies misophonia participants with $t(T, Av) > 10$. **D)** Physiological response to stimuli. Shaded regions represent the 95% bootstrapped confidence intervals. Black overlines show time intervals where differences between responses to trigger and aversive stimuli were statistically significant according to a cluster-based permutation test.

On average, physiological responses to auditory and visual but not mental imagery trials were predictive of stimulus types (Fig. 3C, E). However, these classification results only held for a subset of participants with misophonia (Fig. 3F), similar to results in Figure 2A. These results were primarily driven by responses to triggers in misophonia participants (Fig. 3G, H), particularly those who self-reported being strongly triggered (Fig. 3I). Interestingly, there was greater confusion between triggers and non-aversive stimuli than between triggers and aversive stimuli, suggesting that in some cases, misophonic responses might not have been triggered as expected. Alternatively, if ECG was driving this pattern in the confusion matrices (Fig. 3G-I), the ECG response for non-aversive stimuli being midway between the response for triggers and aversive stimuli (Fig. 2D) could also explain this observation. To test that possibility, we evaluated the relative importance of the classification features. In line

with the results displayed in Figure 2, EMG C was the most important feature, followed by SKT and ECG (Fig. 4). Further, we performed an ablation study where we removed the features derived from ECG and the confusion pattern did not change. Thus, it appears unlikely the difference in ECG response to the different stimulus types is responsible for this observation.

Finally, we assessed the relationship between prediction accuracies for classifying stimulus types (Fig. 3D, F) and misophonia severity scores (Fig. 2C) across participants. As expected, we found a significant correlation between these measures (Fig. 5). However, contrary to our expectations, we did not observe a significant correlation between prediction accuracies and mental imagery scores in the mental imagery condition.

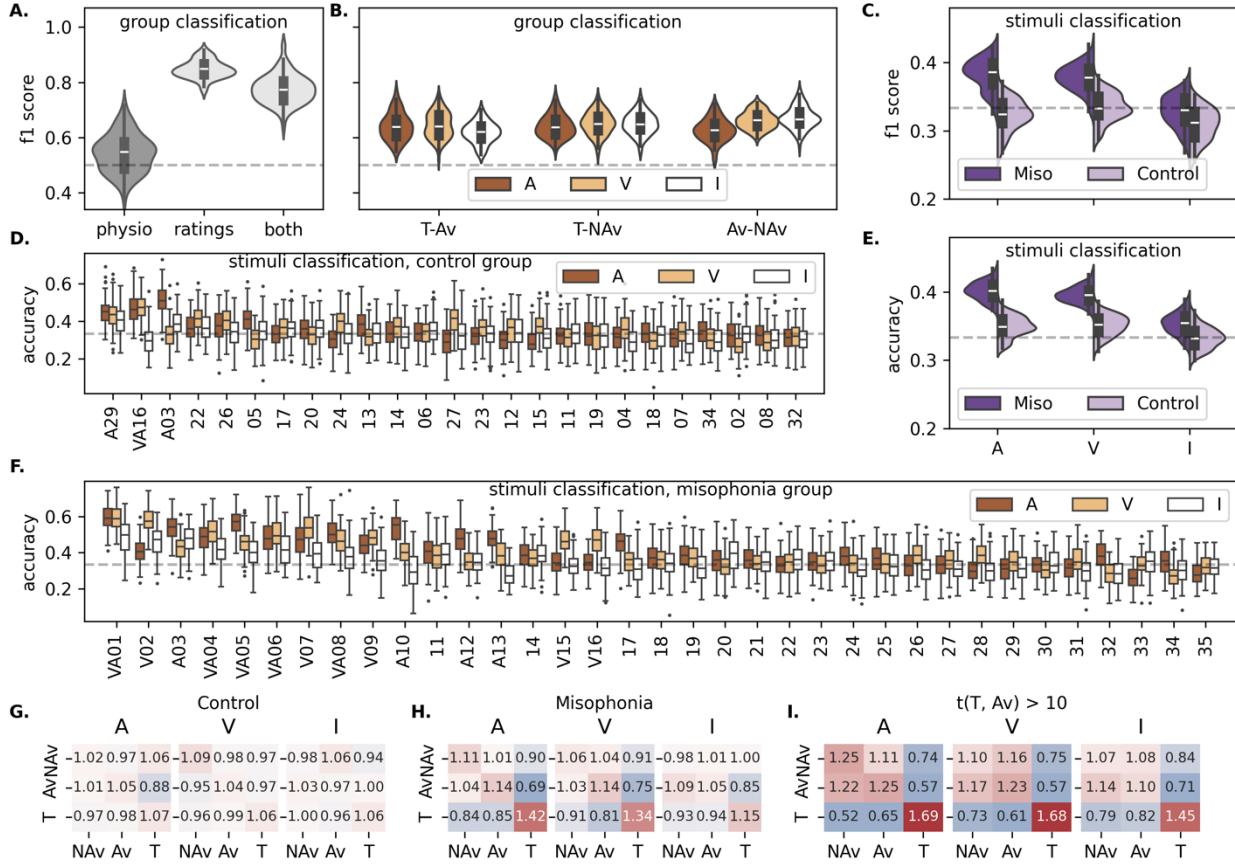


Fig. 3. Predictivity of the physiological response assessed through machine learning. **A)** Prediction (weighted f1 scores) of group membership (misophonia vs. controls) for the average response per sensory modality and stimulus type using physiology data only (i.e., 60 participants; 45 features: 3 stimulus types \times 3 sensory modalities \times 5 physiological signals), self-report ratings only (9 features), or both (54 features). **B)** Prediction (weighted f1 scores) of group membership for each sensory modality (V: visual; A: auditory; I: mental imagery) and stimulus type contrasts (T: trigger; Av: aversive; NAV: non-aversive). **C, E)** Weighted f1 scores (**C**) and accuracies (**E**) for stimulus type prediction per sensory modality. **D, F)** Prediction (weighted f1 scores) for the stimulus types per participant in the control group (**D**) and the misophonia group (**F**). Participants were sorted by decreasing order of accuracy. Participant pairing between groups is indicated using the same numbers. For panels **A-F**, the pale dashed lines indicate the chance level. Participant numbers were prefixed with V, A, or I when the 5th percentile of the bootstrapped accuracy distribution was above chance levels (0.33). **G-I)** Confusion matrices per sensory modality for predicting stimulus types in the control (**G**), the misophonia (**H**) groups, and the misophonia participants with $t(T, Av) > 10$ as defined in Figure 2 (**I**), scaled so that the chance level is 1.0.

EMG C	-3.62	2.30	1.37
EMG Z	-0.06	-0.00	-0.25
EDA	-0.10	0.01	-0.34
SKT	-0.45	0.48	-0.13
ECG	-0.78	0.68	0.03
	A	V	I

Fig. 4. Feature importance for stimulus type prediction by modality. Larger numbers indicate features with a higher influence on the classification.

	All	Control	Misophonia
DVMSQ	0.08 0.02 0.81	-0.08 0.02 0.81	-0.08 0.02 0.81
S-Five	0.03 0.03 0.41	-0.26 0.76 0.57	-0.93 0.40 0.51
BAIS-C	0.68 0.70 0.22	-0.61 0.38 0.98	-0.19 0.18 0.10
BAIS-V	0.97 0.59 0.44	-0.30 0.64 0.65	-0.23 0.19 0.15
VVIQ	-0.46 0.92 0.91	-0.43 0.75 0.46	-0.61 0.90 0.86
	A V I	A V I	A V I

Fig. 5. P-values of Pearson's correlations between stimulus type prediction accuracy (Fig. 3D, F) and both misophonia (Fig. 2C) and mental imagery scores.

4. Discussion

Triggers were robustly identifiable from physiological responses for a relatively small subset of participants with misophonia (Fig. 2A). The absence of significant physiological responses to triggers in some misophonia sufferers may be due to factors such as the failure to replicate the triggering context (e.g., chewing sound in the library) [7], the idiosyncratic nature of the stimuli (e.g., the wrong type of chewing), or insufficient trigger duration. Participants with the highest severity scores were not always the ones for whom triggers were identified, suggesting that we may not have elicited the expected response in all participants. Previous studies used longer stimuli (e.g., 15 s [5]), and as shown in Figure 2D, physiological responses were still building at the end of the 5-second stimulus.

Self-report ratings were sufficient to identify group membership, though not perfectly (Fig. 3A). This could be attributed to several factors: misophonia participants

may not have been triggered by the designated trigger stimuli, self-report measures may have been noisy or unreliable, or the classifier may not have fully captured the complexity of the data. The differences in physiological responses between the stimuli in the misophonia group were more pronounced and driven primarily by trigger stimuli (Fig. 3C-I). Additionally, physiological responses to both aversive and non-aversive non-trigger stimuli differed between individuals with misophonia and those without (Fig. 3B), suggesting that misophonia is associated with a broader intolerance to sound [17-19]. In future work, we plan to explore using autoencoders to extract richer features and apply pre-trained deep-learning models to the time-frequency representation of these physiological signals to improve our ability to identify individual misophonic triggers.

5. Conclusion

Physiological data from multiple recording modalities offer valuable insights into misophonia, highlighting distinct response patterns to trigger and non-trigger stimuli. These physiological responses differ between individuals with misophonia and those without, indicating a clear distinction in how each group reacts to these stimuli. These differences extend beyond sounds to include silent videos and even the mental imagery of sounds. Together, these findings suggest that misophonia may be identifiable through unique physiological patterns, regardless of the stimulus type, highlighting the potential of physiological measures for improving our understanding and diagnosis of misophonia.

Acknowledgment

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

References

- [1] S. E. Swedo *et al.*, "Consensus Definition of Misophonia: A Delphi Study," *Front. Neurosci.*, vol. 16, 2022, Accessed: Mar. 05, 2024. [Online]. Available: <https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2022.841816>
- [2] L. J. Dixon, M. J. Schadegg, H. L. Clark, C. J. Sevier, and S. M. Witcraft, "Prevalence, phenomenology, and impact of misophonia in a nationally representative sample of U.S. adults," *J. Psychopathol. Clin. Sci.*, vol. 133, no. 5, pp. 403–412, Jul. 2024, doi: 10.1037/abn0000904.
- [3] S. Vitoratou, C. Hayes, N. Uglik-Marucha, O. Pearson, T. Graham, and J. Gregory, "Misophonia in the UK: Prevalence and norms from the S-Five in a UK representative sample," *PLOS ONE*, vol. 18, no. 3, p. e0282777, Mar. 2023, doi: 10.1371/journal.pone.0282777.
- [4] A. Schröder *et al.*, "Misophonia is associated with altered brain activity in the auditory cortex and salience network," *Sci. Rep.*, vol. 9, no. 1, p. 7542, May 2019, doi: 10.1038/s41598-019-44084-8.
- [5] S. Kumar *et al.*, "The Brain Basis for Misophonia," *Curr. Biol.*, vol. 27, no. 4, pp. 527–533, Feb. 2017, doi: 10.1016/j.cub.2016.12.048.
- [6] M. Edelstein, D. Brang, R. Rouw, and V. S. Ramachandran, "Misophonia: physiological investigations and case descriptions," *Front. Hum. Neurosci.*, vol. 7, p. 296, Jun. 2013, doi: 10.3389/fnhum.2013.00296.
- [7] M. Siepsiak, S. R. Vrana, A. Rynkiewicz, M. Z. Rosenthal, and W. L. Dragan, "Does context matter in misophonia? A multi-method experimental investigation," *Front. Neurosci.*, vol. 16, Jan. 2023, doi: 10.3389/fnins.2022.880853.
- [8] S. Oh, X. Yang, W. M. Hayes, A. Anderson, D. H. Wedell, and S. V. Shinkareva, "Physiological responses to aversive and non-aversive audiovisual, auditory, and visual stimuli," *Biol. Psychol.*, vol. 195, p. 108994, Jan. 2025, doi: 10.1016/j.biopsycho.2025.108994.
- [9] Z. J. Williams, C. J. Cascio, and T. G. Woynarowski, "Psychometric validation of a brief self-report measure of misophonia symptoms and functional impairment: The duke-vanderbilt misophonia screening questionnaire," *Front. Psychol.*, vol. 13, p. 897901, 2022, doi: 10.3389/fpsyg.2022.897901.
- [10] D. F. Marks, "New directions for mental imagery research," *J. Ment. Imag.*, vol. 19, no. 3–4, pp. 153–167, 1995.
- [11] A. R. Halpern, "Differences in auditory imagery self-report predict neural and behavioral outcomes," *Psychomusicology Music Mind Brain*, vol. 25, no. 1, pp. 37–47, 2015, doi: 10.1037/pmu0000081.
- [12] A. Gramfort *et al.*, "MEG and EEG data analysis with MNE-Python," *Front. Neurosci.*, vol. 7, 2013, doi: 10.3389/fnins.2013.00267.
- [13] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 85, pp. 2825–2830, 2011.
- [14] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A Library for Large Linear Classification," *J Mach Learn Res*, vol. 9, pp. 1871–1874, Jun. 2008.
- [15] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [16] O. Troyanskaya *et al.*, "Missing value estimation methods for DNA microarrays," *Bioinformatics*, vol. 17, no. 6, pp. 520–525, Jun. 2001, doi: 10.1093/bioinformatics/17.6.520.
- [17] N. Andermane, M. Bauer, E. Sohoglu, J. Simner, and J. Ward, "A phenomenological cartography of misophonia and other forms of sound intolerance," *iScience*, vol. 26, no. 4, p. 106299, Apr. 2023, doi: 10.1016/j.isci.2023.106299.
- [18] N. Andermane, M. Bauer, J. Simner, and J. Ward, "A symptom network model of misophonia: From heightened sensory sensitivity to clinical comorbidity," *J. Clin. Psychol.*, vol. 79, no. 10, pp. 2364–2387, 2023, doi: 10.1002/jclp.23552.
- [19] H. A. Hansen, A. B. Leber, and Z. M. Saygin, "What sound sources trigger misophonia? Not just chewing and breathing," *J. Clin. Psychol.*, vol. 77, no. 11, pp. 2609–2625, 2021, doi: 10.1002/jclp.23196.