

# Journal Pre-proof



A Phenomenological Cartography of Misophonia and Other Forms of Sound Intolerance

Nora Andermane, Mathilde Bauer, Ediz Sohoglu, Julia Simner, Jamie Ward

PII: S2589-0042(23)00376-0

DOI: <https://doi.org/10.1016/j.isci.2023.106299>

Reference: ISCI 106299

To appear in: *ISCIENCE*

Received Date: 17 November 2022

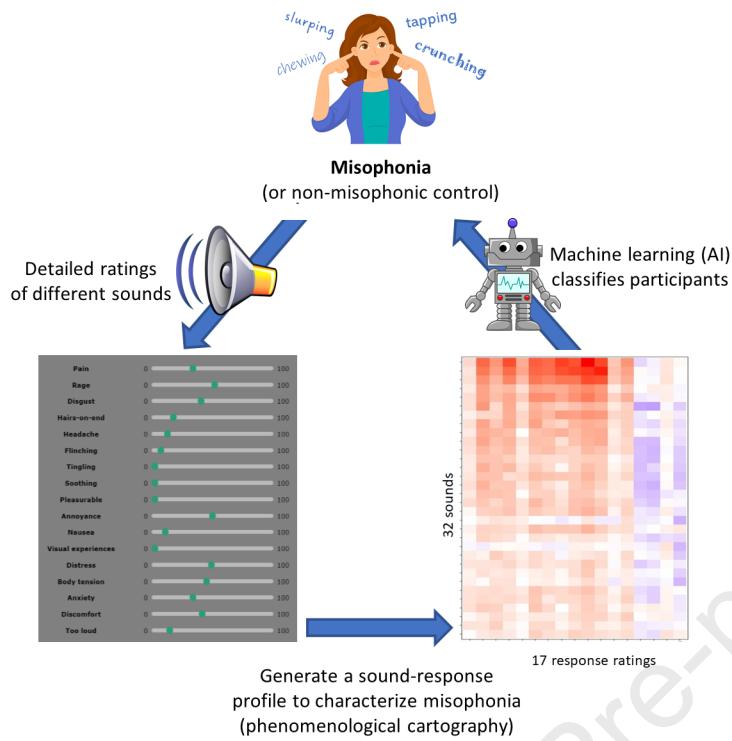
Revised Date: 24 January 2023

Accepted Date: 24 February 2023

Please cite this article as: Andermane, N., Bauer, M., Sohoglu, E., Simner, J., Ward, J., A Phenomenological Cartography of Misophonia and Other Forms of Sound Intolerance, *ISCIENCE* (2023), doi: <https://doi.org/10.1016/j.isci.2023.106299>.

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

© 2023



## A Phenomenological Cartography of Misophonia and Other Forms of Sound Intolerance

Nora Andermane, Mathilde Bauer, Ediz Sohoglu, Julia Simner and Jamie Ward\*

School of Psychology, University of Sussex, Brighton, UK

**\* Lead Contact:**

Prof. Jamie Ward,

School of Psychology,

University of Sussex,

Falmer, Brighton,

BN1 9QH, U.K.

Tel. : +44 (0)1273 876598

E-mail : jamiew@sussex.ac.uk

### **SUMMARY**

People with misophonia have strong aversive reactions to specific ‘trigger’ sounds. Here we challenge this key idea of specificity. Machine learning was used to identify a misophonic profile from a multivariate sound-response pattern. Misophonia could be classified from most sounds (traditional triggers and non-triggers) and, moreover, cross-classification showed that the profile was largely transferable across sounds (rather than idiosyncratic for each sound). By splitting our participants in other ways, we were able to show – using the same approach – a differential diagnostic profile factoring in potential co-morbidities (autism, hyperacusis, ASMR). The broad autism phenotype was classified via aversions to repetitive sounds rather than the eating sounds most easily classified in misophonia. Within misophonia, the presence of hyperacusis and sound-induced pain had widespread effects across all sounds. Overall, we show that misophonia is characterised by a distinctive reaction to most sounds that ultimately becomes most noticeable for a sub-set of those sounds.

**Keywords:** Misophonia, sound intolerance, hyperacusis, autism, ASMR, machine learning

## INTRODUCTION

Misophonia is an extreme aversive reaction to certain sounds (triggers) that can substantially disrupt one's work and social life, as well as negatively impacting on wellbeing<sup>1</sup>. This research addresses the question: what is it that makes a sound misophonic? Is it something to do with the manner (oral v. non-oral) or agent of production (human v. non-human)? Or is it related to some psychoacoustic property of a sound linked to unpleasantness judgments more generally<sup>2</sup>? To what extent are sounds that frequently appear as misophonic triggers also the same sounds that other people (non-misophonics) dislike? Alternatively, it may not be the sounds themselves that best identify misophonia but, instead, a particular kind of response (e.g., rage, anxiety). In this study, we introduce a novel approach that we label 'phenomenological cartography' because we seek to map out what it is that makes a person misophonic using first-person (phenomenological) responses to sounds. Specifically, our map consists of 32 sounds (a mixture of typically described triggers and non-triggers) from which participants give response ratings to 17 descriptors (rage, anxiety, disgust, soothing, etc.). In effect, determining who is misophonic and who is not (from this 17x32 array of numbers) lends itself to machine learning techniques as a form of pattern recognition (where 'pattern' is synonymous with 'multivariate' in statistical terms). Using this approach, we address the questions posed above by determining whether some sounds permit accurate classification of misophonia while others do not, and whether classification depends on the presence of some responses (e.g., rage) more than others. It also leaves open the possibility that there is not a "one-size-fits-all" solution. For example, misophonics may react with 'rage' to the sound of chewing but 'hairs on end' to creaking. This idiosyncratic sound-response pattern is discoverable by our approach.

Misophonic triggers can sometimes appear to be highly specific – such as the sound of one particular member of the family eating – and they can vary from one person to the next. On the face of it, this does not bode well for finding a generalizable misophonic sound-response pattern. But the observation that misophonic triggers are generally described as falling into certain categories of sounds, namely non-vocal oral/nasal sounds (eating, heavy breathing) or other human-made sounds (e.g. tapping) contradicts the idea they are entirely idiosyncratic<sup>3</sup>. Specificity in misophonia could also be conceptualised as a minority of aversive trigger sounds against a backdrop of normal sound tolerance. Evidence in favour of this position comes from neuroimaging studies that show group differences only when misophonics, compared to non-misophonics, listen to trigger sounds (e.g., in the insular cortex) but not for other kinds of sounds<sup>4 5</sup>. However, an absence of a group difference for non-trigger sounds should be interpreted with caution (limited sample sizes, limited sound exemplars). When given sounds to rate for unpleasantness or discomfort, Hansen et al.<sup>6</sup> found that misophonics tended to give higher ratings for all categories of sound (human oral/nasal, human other, non-human) including for ones not traditionally described as triggers (e.g., animal sounds). In summary, there is good evidence that certain sounds are more affected by misophonia than others but the extent to which this reflects a narrow or broad pattern of sound intolerance is unclear.

Outside of misophonia itself, there is also a debate as to how to differentially diagnose different types of atypical sound sensitivity. Williams, He, Cascio and Woynaroski<sup>7</sup> argued that different forms

of sound intolerance should have different kinds of specific reactions not well captured by umbrella terms like ‘unpleasantness’. Specifically, they argue that misophonia should elicit anger, extreme annoyance, and disgust; hyperacusis should be linked to loudness and pain; and phonophobia should be linked to fear and anxiety. They argue that decreased sound tolerance linked to the autism spectrum should have features of all three of these (i.e., be much broader in response categories). Others have argued that misophonia is primarily characterised by physical reactions such as muscle tension<sup>8</sup>. Finally, a non-clinical trait linked to sound over-responsivity is ASMR, autonomous sensory meridian response e.g.<sup>9</sup> for which a more accurate term may be audio-visual elicited somatosensation<sup>10</sup>. ASMR is typically described as a pleasurable experience triggered by sounds linked to physical reactions such as tingling (notably to the hair/head). ASMR triggering sounds are often described as whispering, popping and rustling sounds. It has even been suggested that ASMR and misophonia may tend to co-occur<sup>11-13</sup>. That is, misophonia may be linked to atypically intense experiences to sounds that, at least in some instances, have a positive valence<sup>8</sup>.

In the present pre-registered study, we aim to document the sound-response profile of misophonia using phenomenological cartography. We use machine learning to classify each and every sound and also perform cross-classification: to determine whether a misophonic response to, e.g., the sound of apple crunch can be used to predict misophonia status from responses to a completely different sound. Successful cross-classification would indicate a general response profile common for all sounds. Natural sounds are selected from the same three categories as Hansen et al.<sup>6</sup>, i.e., including sounds not normally considered as triggers, plus a fourth category which is made up of ‘scrambled’ versions of misophonic triggers (such that they have similar acoustic properties but are unrecognisable). As this study is part of a wider project<sup>14</sup>, we are also in the fortunate position of having a set of detailed measures on other relevant individual differences that speak to the question of a differential diagnosis. For the misophonics, they were asked (in a separate session) about hyperacusis and about pain responses to sounds. Hyperacusis is typically defined by the perception that sounds feel subjectively too loud, sometimes causing pain in the ears or head<sup>15,16</sup>. For the non-misophonics, we split the data according to being high/low in autistic traits<sup>17</sup>, high/low sensory sensitivity across all senses<sup>18</sup>, and high/low interoceptive awareness<sup>19</sup>. Our goal here is to ascertain whether any misophonia profile is specific to misophonia, rather than reflecting a co-morbid trait also linked to atypical sound tolerances. Finally, in a separate study, we recruit a group of people with ASMR and add a set of ASMR trigger sounds which are presented alongside the same set of sounds given to misophonics.

## RESULTS

The method and analysis plan was pre-registered (<https://osf.io/27fgv/>) and any deviations from this are explicitly noted.

### A Phenomenological Cartography of Misophonia

For each sound, responses are measured using a 0-100 visual analog scale to descriptors such as rage and anxiety. Figure 1 shows the 32 sounds x 17 responses displayed as a heatmap, with colour

denoting the magnitude of the group difference for misophonics versus non-misophonics (Cohen's d). A classifier (a machine learning algorithm) was developed for each of the 32 sounds in which it attempts to predict group status from the 17 responses, taking different training versus test resamples of the data. AUC accuracy (area-under-curve, where 0.5 is chance and 1 is perfect classification) is indicated for each row, reflecting the ability of the classifier to predict group status. For visualisation purposes, the sounds are ranked, top to bottom, in terms of AUC, and the responses are organised so that frequently co-occurring ones are nearby (hierarchical clustering based on Pearson's correlation; see Figure S6 for the corresponding dendrogram).

Thirty sounds could discriminate misophonics from controls above chance (where significance above chance was calculated via randomly permuted group labels), and two were at chance (scrambled breathing and scrambled coughing). All 30 significant results are retained after FDR correction for multiple comparisons ( $p < .05$ ). The most discriminating sounds were well-documented misophonic triggers (crunching chips, apple crunch). However, the 32 sounds do not cluster, in a strict way, according to our initial four categories (human oral/nasal, human actions, non-human, scrambled). Instead, the overall finding is of a broad pattern of atypical responsiveness to most sounds including to some of the unrecognizable scrambled sounds. Training a single classifier across the whole 32x17 (N=544) multivariate pattern results in a diagnostic accuracy for misophonia of AUC = 0.925 (specificity = 0.865, sensitivity = 0.802) which compares favourably against questionnaire approaches<sup>20,21</sup>.

With regards to which features are important for classification, the VSURF algorithm Variable Selection using Random Forests,<sup>22</sup> determines the most important features for successful performance of the classifier. Summing across all 32 classifiers the most important features were (in descending order): body tension (20), annoyance (19), anxiety (16), visual experiences (16), rage (15), discomfort (15), too loud (14), soothing (13), pleasurable (10), flinching (10), distress (9), disgust (8), pain (4), tingling (2), hairs-on-end (1), and nausea (1). Three of these features tended to yield higher responses in controls relative to misophonics (pleasurable, soothing, visual experiences) – as depicted in blue. A simple averaging of response effect sizes, down each column (ignoring sign), produces a similar rank order: body tension, rage, anxiety, distress, annoyance, discomfort, flinching, too loud, disgust, pain, soothing, hairs on end, visual experiences, pleasurable, headache, nausea, tingling.

#### FIGURE 1 ABOUT HERE

#### Comparisons between Sounds: Cross-classification

Cross-classification is a way of testing if there is a common response profile across sounds. That is, we take the classifier for one sound (e.g., the 'apple crunch' classifier) and apply it to the set of responses evoked by every other sound (the ratings of rage, annoyance and so on elicited by sounds of 'chewing gum', 'dog drinking', etc.). We then repeated this for each of the 32 sounds in our stimulus set, yielding a 32x32 matrix of cross-classification accuracies. Figure 2 visualizes this matrix, where the on-diagonal cells represent classification performance within sounds (i.e. the AUC values already noted in Figure 1) and the off-diagonal cells represent cross-classification between sounds (again expressed as AUC). For visualization, we group the 32 sounds into the four original categories. The rows are

different classifiers and the columns are different response profiles to enter into the classifier. Overall, there is a high degree of cross-classification: 58% of cross-classifications are significant at  $p < .05$  and 50.1% are retained after FDR correction for multiple comparisons. Responses to human oral/nasal sounds are the easiest to classify (shown here in the left eight columns). But this is generally true across most classifiers. For example, when we train a classifier to predict misophonia status from scrambled breathing sounds it has AUC = 0.541 but when we test the same classifier on a new set of responses to chewing gum it has AUC = 0.646. That is, the response pattern from this meaningless sound acts as a good predictor (/a better predictor) of a trigger sound. The fact that the classification accuracy appears to go up in this example is because the group differences are more extreme for chewing gum. But the fact that cross-classification occurs at all is indicative of a common response pattern across a wide range of sounds.

#### FIGURE 2 ABOUT HERE

#### Comparisons between Sounds: Why are some sounds rated highly?

One possibility for why some sounds generate consistently high ratings across descriptors is that they contain common acoustic features. To test this, for each of our 32 sounds, we measured the energy in the psychoacoustic space of 2500-5500, Hz and 1-16 Hz temporal modulations. This has previously been shown to predict unpleasantness ratings for natural sounds<sup>2</sup>. Across our set of 32 sounds, we correlated these values with the mean ratings across participants (separately for each participant group and for distress, rage, etc.). There were four significant correlations when considering all 32 sounds together although none survived FDR correction (headache/controls  $r = 0.345$ , tingling/misophonia  $r = 0.471$ , too loud/controls  $r = 0.471$ , visual experiences/misophonia  $r = 0.373$ ). However, we suggest it is not meaningful to group all the sounds together. Post hoc we noted many large (and significant) correlations for our unrecognizable scrambled sounds, but far smaller correlations for the natural sounds. This was true both for the misophonics and controls (see Figure 3). Given the limited number of sounds in our study, we conducted a reanalysis of the N=125 natural sounds from Hansen et al.<sup>6</sup>. Here we use their published mean ratings of discomfort (which is also one of the 17 responses in our study) and conducted our own psychoacoustic analysis of their sound files (up to the first 30s). There was no significant correlation between mean discomfort ratings from either misophonics ( $r = -0.028$ ) or controls ( $r = 0.004$ ) and the energy in the region of interest (2500-5500 Hz and 1-16 Hz temporal modulations). In summary, we conclude that whilst this psychoacoustic feature has some relevance for aversive responses to sounds (as indicated by the large correlations for scrambled stimuli), it tends to be masked by other factors when sounds are easily recognizable. This occurs irrespective of misophonia status (although further research is warranted given the small number of scrambled sounds used here).

#### FIGURE 3 ABOUT HERE

More generally, one can ask the question whether the sounds which induce annoyance (rage, etc.) are the same or different across groups. Put simply, if we were to rank the 32 sounds according to their mean ratings of annoyance for misophonics then again for controls, would we find these rankings to be similar? Across all 17 features, there is a large and statistically significant correlation between the cross-group ranking of the sounds (see Figure 4). That is, although misophonics tend to rate sounds more negatively (in absolute terms) the relative difference between the sounds (i.e., ranking them from least to worst) is largely preserved between both groups.

#### FIGURE 4 ABOUT HERE

#### Heterogeneity within Misophonics

This rich dataset offers opportunities for exploring individual differences amongst misophonics. As a starting point, one can identify misophonics who are easy or hard to classify according to the number of sounds for which they were correctly classified as having misophonia (a score from 0 to 32). For the misophonics, the mean number of sounds that led to an accurate group classification was 22.17 and S.D. was 5.11 (range = 5 to 32). These differences amongst misophonics in classification performance were significantly correlated with self-reported severity and breadth of misophonia from data collected in a previous session using the Sussex Misophonia Scale, SMS<sup>21</sup>. The Pearson's correlations between classification accuracy and the number of triggers or total questionnaire score were  $r = .381$  and  $.417$  respectively ( $p < .001$ ), as shown in Figure 5. Supplemental Figure S13 shows heatmaps for categorical differences in severity, comparing the transition from non-misophonia to moderate misophonia, and from moderate to severe misophonia (based on the three groups identified in our prior research<sup>14</sup>). These results show that our machine learning approach is tracking real-world severity and, moreover, that misophonia severity is linked to a less specific profile (i.e., extending to more sounds).

#### FIGURE 5 ABOUT HERE

#### Differential Diagnostic Profiles: Within Misophonia

As our participants had been asked about the presence of other traits and co-morbidities, we can determine whether the pattern observed for misophonia is specific to misophonia. Effect sizes (Cohen's d) were computed for different splits of the dataset. They are displayed as heatmaps for visualization purposes, and we trained a classifier (as before) to determine which sounds predict the various group memberships. The potential presence of hyperacusis within the misophonics was ascertained in two ways: either by directly asking them (based on<sup>15</sup>) or through a set of questions included within the SMS that ask about pain. These two different approaches yield different results, as shown in Figure 6, which depicts differences in responses as a function of self-reported hyperacusis (present vs absent) and pain (present vs absent). The results emphasise the importance of different responses (which are visualised as prominent columns), rather than different sounds. Those misophonics who agree to the hyperacusis question went on to give higher responses for 'body tension'

(across most sounds) whereas those who agreed to the ‘pain’ questions on the SMS went on to give higher responses for headache and pain (across most sounds) – noting that these latter two features are not the strongest indicators of misophonia per se (see Figure 1 for example). The association between pain measures is not entirely trivial given that the measures were given months apart and differed substantially in method (trait-based questionnaire versus state-based responses to specific sounds). For classifiers trained to predict high/low pain for misophonics as reported in the SMS questionnaire, these could predict group membership above chance ( $p < .05$ ) for 20 sounds with 12 remaining significant after FDR correction: sniffling (AUC=0.672), pen clicking (AUC=0.662), breathing (AUC=0.658), rusty swing (AUC=0.641), crunching chips (AUC=0.639), chewing gum (AUC=0.638), typing (AUC=0.634), nails on chalkboard (AUC=0.628), clock ticking (AUC=0.625), basketball (AUC=0.625), chopping vegetables (AUC=0.620), and finger tapping (AUC=0.618). For classifiers trained to predict self-reported hyperacusis, these could predict group membership above chance ( $p < .05$ ) for 13 sounds with six remaining significant after FDR correction: scrambled typing (AUC=1.000), finger tapping (AUC=0.649), basketball (AUC=0.629), scrambled coughing (AUC=0.624), scrambled apple crunch (AUC=0.622), and typing (AUC=0.613). In summary, we show that comorbid traits of hyperacusis and sound-induced pain have a different sound-response profile to misophonia and that the phenomenological cartography approach is sensitive to different forms of sound intolerance.

#### FIGURE 6 ABOUT HERE

#### Differential Diagnostic Profiles: Within Non-misophonics

With regards to individual differences amongst the non-misophonic controls, the effect sizes are generally less extreme. High versus low autistic traits (measured by the AQ) has some similarities with misophonia in that the four responses depicted on the right of the heatmap in Figure 7 showed the same reversed ‘blue’ profile as misophonics (pleasurable, soothing, tingling, visual experiences). But, unlike misophonia, the top-to-bottom ordering of effect sizes was not apparent and – instead – there was a pervasive low-level negativity towards most sounds. The most discriminating sounds between high v. low AQ were in descending order: scrambled basketball (AUC=0.692), typing (AUC=0.690), finger tapping (AUC=0.637), basketball (AUC=0.631), dog drinking (AUC=0.612), and rusty swing (AUC=0.601). These were all significant with the top four surviving FDR correction.

Neither the data splits of high v. low sensory sensitivity (using the GSQ) or high v. low interoceptive awareness (using the MAIA) produced any significant results that survived FDR correction (see Figures S14 and S15).

#### FIGURE 7 ABOUT HERE

#### Comparison of Misophonia and ASMR Responders

A separate set of  $N = 254$  participants were recruited and given questionnaires relating to misophonia (SMS) <sup>21</sup> and ASMR (ASMR-15) <sup>23</sup>. In addition, they completed the phenomenological

cartography study above with the original 32 sounds plus the addition of a further set of 8 ASMR triggers (e.g., whispering). There was no significant correlation between the ASMR-15 and the SMS scores ( $r = .113$ ,  $p = 0.072$ ). We applied the aggregated classifier from the previous sample (pooled from 32 sounds and 17 responses) to predict the misophonia status of these new participants. The probability of being classified as misophonic (from the machine learning algorithm) was correlated with the SMS scores ( $r = .399$ ,  $p < 0.001$ ) but not the ASMR-15 scores ( $r = -.089$ ,  $p = 0.159$ ). The differences in correlation were significant ( $t(251) = 6.368$ ,  $p < .001$ ).

Excluding misophonics (i.e., those with SMS scores greater than the diagnostic threshold of 50.5), we created a separate phenomenological cartography heatmap contrasting high and low ASMR groups, by applying a mean ASMR-15 value of 2.5 (1-5 scale) to divide the groups. This gives a mean ASMR-15 score of 3.75 in the high group which is comparable with other studies of ASMR responders (3.49 in <sup>24</sup>, and 3.73 in <sup>23</sup>). The ASMR heatmap for the main set of 32 sounds is shown in Figure 8. The strongest group differences are found for the responses of soothing, tingling and pleasurable, and were most strongly elicited by the scrambled sounds. ASMR status could be classified above chance from 21 sounds (after FDR correction) suggesting that ASMR, like misophonia, is not limited to its traditionally conceptualised triggers. The eight ASMR triggers shows the same basic pattern (albeit with far larger effect sizes). These sounds also tended to be rated lower in annoyance and discomfort by ASMR responders relative to non-responders. Figure S16 shows the corresponding data for misophonics in this sample (with lower effect sizes and a pattern more similar to that documented previously for misophonia e.g., Figure 1), and those reporting both ASMR and misophonia (who tend to resemble the ASMR profile). In summary, we show that misophonia and ASMR have distinctive profiles with very little in common.

#### FIGURE 8 ABOUT HERE

## DISCUSSION

There are significant individual differences in the ability to tolerate (or enjoy) certain sounds which, in extreme situations, impacts on quality of life enough to require intervention. We developed an approach, which we term phenomenological cartography, to map out the sound-response profile of misophonia. We use this as a predictive tool by applying machine learning to classify participants. We demonstrate that, despite heterogeneity, there is a distinctive pattern linked to misophonia that extends across most sounds (but is clearly more extreme for certain sounds). There are individual differences within this profile: more severe misophonics are easier to classify, and the presence of sound-induced pain and/or hyperacusis within the misophonics can be considered a separate 'layer' superimposed on the map (for further discussion of hyperacusis subtypes <sup>25</sup>). We further show that the misophonic sound-response pattern is different to the sound-response profile linked to the broad autism phenotype or ASMR. The former is linked to repetitive sounds (e.g., typing, basketball dribbling) and the latter to scrambled sounds, whispers, etc. In effect, phenomenological cartography enables us to dissect a complex set of subjective sound-response relationships into distinct cognitive profiles. This is important for driving scientific theory and for real-world practice (differential diagnosis and possibly treatment).

Misophonia is defined as a disorder of decreased tolerance to specific sounds<sup>1</sup>, but our research challenges us to rethink the idea of specificity. Specificity could be taken to mean that a minority of sounds evoke misophonic responses against a backdrop of otherwise normal sound tolerances. Alternatively, specificity could be taken to mean that misophonic sounds are highly idiosyncratic or personal. Our data is incompatible with these interpretations but supports, instead, a weaker conclusion that some sounds are far worse for people with misophonia than others. Importantly, this sits against a backdrop of generalized sound intolerance and a high level of agreement (across individuals and groups) about which sounds are aversive. Similarly, there are a broad range of responses linked to misophonia although some (e.g., body tension, rage) were found to be more important than others (e.g., nausea, hairs on end). If specificity were the hallmark of misophonia then it is also puzzling to find an inverse relationship between specificity and severity: more severe misophonics have a less specific profile. In effect there are two key properties of misophonia that need to be explained: why there is a pervasive pattern of sound intolerances (generality) and why some sounds are worse than others (weak specificity). Whether a single explanation can account for both observations is unclear and is discussed further below.

It is conceivable, at least in theory, that specificity is an emergent property of generality such that there is a ‘gain’ or amplification of the normal hierarchy of sound intolerances<sup>26</sup> – i.e., the degree of worsening in misophonia may be related to how aversive they are to begin with. We provide evidence that misophonics and controls do rank sounds similarly (across 17 rating scales) which speaks to this idea. However, there are important exceptions that suggest a more complex picture. To give two examples. ‘Rusty [creaking] swing’ and ‘nails on chalkboard’ are strongly disliked by both groups but the relative difference between the groups is marginal (as noted elsewhere<sup>2</sup>). On the other hand, crunching chips are far more misophonic than the sound of coughing (despite the latter having clear negative connotations). It is likely that the degree of aversion to sounds reflect a complex evaluation of multiple sources of information including its meaning (semantic content, social context, etc.) and psychoacoustic properties. Changing the meaning of a sound (e.g., by presenting it with an incongruent video<sup>27</sup>) or making it harder to identify (e.g., presenting it against noise<sup>28</sup>) reduces the gap between misophonics and controls although it does not necessarily eliminate it (as shown also for our scrambled sounds). But whether a single dimension can predict the relative ranking of sounds (misophonic > control) remains to be shown.

Here we also tested whether a psychoacoustic feature linked to sound unpleasantness (frequencies between 2500-5500, Hz and 1-16 Hz temporal modulations<sup>2</sup>) predicts ratings from misophonics and non-misophonics. Our own dataset has the advantage of multiple rating scales but the disadvantage of a low number of sounds. But we also analysed the dataset of Hansen et al.<sup>6</sup> that is complementary to our own (many more sounds albeit rated in a more limited way). In neither dataset do we find evidence of a relationship between this psychoacoustic feature and ratings of natural sounds. However, we do find evidence of a relationship with our unrecognizable scrambled sounds (and so we speculate that many of the natural sounds in the original study by Kumar et al.<sup>2</sup>, may also have been hard to recognize). As we only focussed on this one psychoacoustic feature it remains a possibility that other (unknown) psychoacoustic properties are relevant to misophonia. It may also be the case that

the psychoacoustic properties of sounds are more relevant for autism or ASMR than they are for misophonia. In ASMR it was the unrecognizable sounds that were most potent which suggests a reliance on sound properties over meaning. The repetitive sounds found to be predictive of autism extended to both natural and scrambled sounds, pointing to the same conclusion.

Misophonia has been found to be comorbid with autism<sup>29</sup>, hyperacusis<sup>14,30</sup>, and ASMR<sup>11-13</sup>. It is important to speculate how these comorbidities can co-occur if they have different sound-response profiles. One possibility is that the profiles are layered on top of each other (i.e., additive in statistical terms). For example, someone with misophonia and autism may have both particularly strong responses for misophonic triggers (chewing, etc.) and those linked of autism (tapping, etc.) against a backdrop of general sound intolerance that acts as a common denominator to both conditions. Hyperacusis could be understood in a similar way, albeit manifesting itself more in terms of a characteristic response profile (e.g., pain) than sound specificity. That is, it is conceivable that sound intolerances in misophonia, hyperacusis, and autism all stem from a general sensory hyper-sensitivity that, whilst strongest for sounds, is by no means limited to it<sup>31-33</sup>. In effect, generalized sensory hyper-sensitivity may be a common seed from which somewhat different profiles emerge during development. A comorbidity between misophonia and ASMR is harder to understand as layering of different profiles because they would tend to cancel each other out (e.g., ASMR responders report high pleasure and soothing, and misophonics report the opposite). Indeed, our misophonia classifier does not predict ASMR-15 scores (the correlation is, if anything, negative). The self-report questionnaires used in our study hint at a weak positive relationship. This may reflect superficial similarity (sound over-responsiveness) not borne out by detailed assessment. ASMR and misophonia may also arguably differ in terms of whether the response is primarily somatosensory or affective/interoceptive respectively<sup>10</sup>.

One surprising difference between our groups was that misophonics reported less ‘visual experiences’. Further research is needed to understand how participants interpreted this term, but we speculate that it is related to mental imagery (e.g., because visual experience ratings here were uniformly low for scrambled sounds which can’t be easily imagined). By contrast, such a pattern would not be expected for sound-vision synesthesia, where visual experiences are elicited by all sounds, and where these experiences are very rare<sup>34</sup>. An fMRI study also found that misophonics have less activity (compared to non-misophonics) in visual ventral stream when listening to misophonic triggers<sup>5</sup>, which is consistent with our results. Further research is needed to understand whether this is a general difference in the capacity to generate visual images, or if it is context specific. It may be that when listening to sounds, misophonics divert attention inwards (to the reaction in their body) whereas other people imagine the source of the sound. If so, this could potentially be developed as a target of treatment for misophonia, in the form of strategic reorienting.

In sum, the novel approach of phenomenological cartography has revealed important insights into the likely mechanisms of misophonia (it is linked to broad sound intolerance), the differential diagnoses of other atypical sound sensitivities (e.g., requiring minimal translation to be used cross-culturally), and possible treatments for misophonia (visualizing the source).

### **Limitations of the Study**

Machine learning is commonly criticised for lack of interpretability of results (a ‘black box’ approach) and the risk of over-fitting (inflated performance due to, e.g., detecting confounds). Both are valid limitations of the present study, although we made attempts to mitigate them. Interpretability was aided by performing more conventional analyses in parallel (e.g., heatmaps derived from mass univariate analyses) and separate train and test resamples guards against over-fitting. Other choices of machine learning algorithms were not compared. Finally, whilst we took advantage of having rich phenotypic information about our participants (enabling us to split the dataset in different ways) further research is needed to validate differential diagnoses on clinically defined autism and hyperacusis. The generalizability to other demographics also needs to be established, including cross-culturally.

### **Acknowledgments**

This research was financially supported by the Misophonia Research Fund of the Ream Foundation. We thank Heather Hansen for sharing the sound files from their published study, and also thank the late Prof. David Baguley for his advice on the study design and measurement of sound intolerances including hyperacusis. Dodi Swan-Capper assisted with the piloting and data collection for the ASMR sample. The graphical abstract contains copyright-free icons from vecteezy.com and pixabay.com.

### **Author Contributions**

The study was designed by NA, JW and JS. ES advised on sound scrambling and the psychoacoustic analysis of sounds. NA was responsible for setting up the study and for data collection. NA and MB conducted data pre-processing, and machine learning analysis were conducted by JW. JW wrote the manuscript, and all authors contributed to editing it.

### **Declaration of Interests**

No conflicts of interest are declared.

### **Inclusivity and Diversity**

We support inclusive, diverse, and equitable conduct of research.

### **Main Figure Titles and Legends**

Figure 1: A heatmap depicting misophonic > control effect sizes (Cohen’s d) for 32 sounds and 17 response features. The sounds are ranked according to the ability of a machine learning classifier to predict group membership (AUC). \* p < .05 significance relative to chance classification. See also Figure S2.

Figure 2: A heatmap depicting cross-classification accuracy (AUC) when a classifier trained on responses to one sound is used to predict group membership from a pattern of responses to another sound (note that the on-diagonal values are not cross-classifications). The grid lines show divisions

between the four sound categories (human oral/nasal, human actions, non-human, and scrambled). The percentages within each grid cell refer to the number of cross-classifications that achieved significance (at  $p < .05$  FDR corrected).

Figure 3: Top: Spearman's correlations between the psychoacoustic region-of-interest identified by Kumar et al.<sup>2</sup> and phenomenological ratings to the same sounds by misophonics and controls considering scrambled and real sounds (left and right). Note, the critical values for alpha = 0.05 are 0.643 ( $n = 8$ , i.e. scrambled sounds) and 0.344 ( $n = 24$ , i.e. natural sounds). \*  $p < .05$

Figure 4. As an example, mean ratings for 'annoyance' ranked from most to least annoying (according to the misophonic mean rating), with the mean correlation between ranks being 0.912. Equivalent plots for all 17 descriptors are provided in Supplemental Figures S8-S12. The right panel shows Spearman's correlations between ratings to the same sounds across groups. The critical value for alpha = 0.05 is rho  $\geq 0.296$  (for  $N = 32$  sounds).

Figure 5: The ability to classify misophonics based on their response pattern to 32 different sounds (x-axis) correlates with the number of misophonic triggers reported (top figure) and the severity of the misophonia (bottom). Each point is an individual misophonic participant.

Figure 6: Heatmaps (showing Cohen's d effect sizes) for 32 sounds (y-axis) and 17 responses (x-axis) for within misophonia comparisons, contrasting presence versus absence of hyperacusis (top) and high versus low scores on the SMS pain factor (bottom).

Figure 7: Heatmaps (showing Cohen's d effect sizes) for 32 sounds (y-axis) and 17 responses (x-axis) for comparisons within non-misophonics, contrasting high versus low AQ scorers.

Figure 8: Heatmaps (showing Cohen's d effect sizes) for (top) 32 sounds (y-axis) and 17 responses (x-axis) and (bottom) the eight ASMR triggers, contrasting ASMR responders and non-responders (excluding misophonics).

## STAR METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Deposited data</b>		
Raw and analyzed data	This paper	<a href="https://osf.io/27fgv/">https://osf.io/27fgv/</a>
<b>Software and algorithms</b>		

Time-domain scrambling of audio signals in Matlab	Ellis <sup>35</sup>	<a href="https://www.ee.columbia.edu/~dpwe/resources/matlab/scramble/#1">https://www.ee.columbia.edu/~dpwe/resources/matlab/scramble/#1</a>
NSL (Neural Systems Laboratory) auditory-cortical Matlab toolbox	Chi, Ru and Shamma <sup>36</sup>	<a href="http://nsl.isr.umd.edu/downloads.html">http://nsl.isr.umd.edu/downloads.html</a>
VSURF, Variable Selection using Random Forests, <sup>22</sup>	Genuer et al. <sup>22</sup>	<a href="https://journal.r-project.org/archive/2015/RJ-2015-018/index.html">https://journal.r-project.org/archive/2015/RJ-2015-018/index.html</a>
<b>Other</b>		
Pre-registered methods and analysis plan	This paper	<a href="https://osf.io/27fgv/">https://osf.io/27fgv/</a>
Sound stimuli	This paper	<a href="https://osf.io/27fgv/">https://osf.io/27fgv/</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and materials may be directed to and will be fulfilled by the lead contact, Prof. Jamie Ward (jamiew@sussex.ac.uk).

### Materials availability

- Sounds used and created in this research are available for download and re-use (without permission) from <https://osf.io/27fgv/>

### Data and code availability

- All raw data has been deposited at Open Science Framework (<https://osf.io/27fgv/>) and is publicly available as of the date of publication.
- All original code has been deposited at Open Science Framework (<https://osf.io/27fgv/>) and is publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study was approved by the Cross-Schools Science and Technology Research Governance and Ethics Committee of the University of Sussex, UK.

### Sample comparing Misophonia and Non-misophonic Controls

A total of 418 participants completed the task and were divided into two groups: a non-misophonic control group (N = 196; mean age = 33.08, S.D. = 14.29; gender = 140 female, 55 male, 1 non-binary) and a misophonic group (N = 222; mean age = 41.56, S.D. = 14.32; gender = 175 female, 42 male, 5 non-binary). Misophonia is equally prevalent across sexes<sup>37</sup> and the over-representation of women in our samples is attributed to a recruitment bias. This sample size is sufficient to detect Cohen's d of 0.3 and above (with power = 0.85) in univariate analyses. Details of their recruitment and

assignment into groups are documented elsewhere<sup>14</sup>, but briefly restated here. The Sussex Misophonia Scale<sup>21</sup>, rather than recruitment source, was used to determine group status. Here we used a clustering method based on the five factors of this scale to identify three groups: non-misophonic, moderate misophonia, and severe misophonia. The latter two groups are collapsed in the present study when performing a binary classification, and we note that the binary group division closely corresponds to the published cut-off value of > 50.5 using this measure<sup>21</sup>. The mean total SMS score of non-misophonics was 9.026 (S.D. = 9.006) and for misophonics it was 82.112 (S.D. = 22.767).

### Sample comparing Misophonia and ASMR Responders

Participants were recruited from the psychology undergraduate population at the University of Sussex, but with some additional participants recruited from our misophonia database and an ASMR discussion forum (on reddit) to ensure we received sufficient high responses on these measures. Misophonia and ASMR tendencies were determined by the relevant measures rather than self-declaration or recruitment source. A total of N=254 participants completed all relevant parts of the study (sound ratings, ASMR-15, SMS) and took a sufficiently long time to complete the sound task (> 10 minutes). The mean age was 23.224 years (S.D. = 8.165) with 202 women, 45 men, and 7 non-binary.

## METHOD DETAILS

### Sound stimuli

The final set of sound stimuli consisted of a core set of 32 sounds, together with an additional set of 8 sounds (ASMR triggers) which were given to a second sample.

The core set of 32 sounds selected for this task were downloaded from a free sound database (<https://freesound.org/>) and spliced or concatenated to be 15 seconds long. Their selection was motivated by two previous studies investigating the reactions of misophonics and controls to natural sounds from different categories, and not limited to known misophonic triggers<sup>6,38</sup>. Our sound clips were eight everyday sounds from each of three categories; these are human-made sounds originating from the nose or mouth (i.e., chewing apple, eating crisps, chewing gum, lip smacking, coughing, snoring, breathing, sniffling), other human-made sounds (i.e., playing basketball, chopping vegetables, clicking a pen, finger tapping, nails on chalkboard, swinging on swing set, typing, walking in heels), non-human sounds (i.e., birds singing, clock ticking, crow cawing, dog drinking water, frog croaking, printer, rainfall, wind chimes).

We also present a new category that are used as control sounds; these sounds were created by using four of the human oral/nasal and four of the other human sounds (i.e., apple eating, breathing, coughing, sniffling, basketball, chopping vegetables, finger tapping, typing), and scrambling them in time. Scrambling was done by selecting 15ms snippets from the sound and randomly re-arranging these within 400ms from the original snippet location in the sound file using existing MATLAB (Mathworks Inc., Natick, MA) scripts<sup>35</sup>. This method of scrambling preserves the original spectral content of a sound whilst removing local temporal structure, thus making it less recognisable whilst retaining its spectral similarity to the original sound.

The final selection of the 32 sounds was validated in two earlier pilot studies (N=28 and N=22 recruited from students at the University of Sussex), with natural sounds and scrambled sounds presented to different groups (i.e., so they did not prime each other). The pilot included several additional sounds that were subsequently discarded. The primary aim was to check that participants could recognise the natural sounds, and that the scrambled control sounds were not recognisable. Loudness and discomfort ratings were also collected for each sound using a 5-point Likert scale (none, a little, some, lots, extreme). Real sounds were on average highly recognisable (Mean recognisability = 87.4%, SD = 15.7%), whereas the scrambled sounds were on the whole not-recognisable (7.4% recognisability, SD = 12.4%). The loudness of 7 real sounds were adjusted by scaling intensity in Praat<sup>39</sup> based on the average subjective loudness rating from the pilot data, and the corresponding scrambled sounds were re-scrambled from the loudness-adjusted originals.

The eight ASMR triggers were selected from an initial pool of 16 after piloting them on a separate group (N=24; mean age = 26.5 years, SD = 5.65; 15 female, 8 male, 1 non-binary). The pilot sample were recruited from an ASMR reddit forum, all self-identified as having ASMR, and the average score on the ASMR-15 was 3.86 (S.D. = 0.43). The sound clips for the pilot lasted one minute. Participants were asked after each sound whether they experienced ASMR. They were also asked to press the spacebar when they first began to experience ASMR (start time) and 'enter' when their ASMR peaked (peak time). The eight most reliable ASMR triggers, eliciting a response in at least half of the pilot sample, were selected. In the selected stimuli, the average start time for onset of ASMR was 21.79s (S.D. = 5.09, range=17-30) and the average peak time was 35.30 (S.D. = 6.06, range=26-46). The original sound files were trimmed to 30s each for the main study (i.e., after ASMR onset and close to peak). A full list of triggers, and the pilot data, are included in Table S1.

### Procedure

Participants were required to use a computer screen, with headphones or a computer speaker. They were initially played a musical clip and were asked to adjust the loudness to be as loud as possible without causing any discomfort. They were instructed to maintain this loudness for the duration of the task. Participants listened to 32 sound clips (or 40 in the follow-up study) once or repeatedly if necessary. These were played in a random order. Participants then rated their reactions to each of these sounds. There were 17 descriptors used for the ratings and these were rated on a visual analog scale (ranging from 0 to 100) with endpoints marked as 'none' and 'extreme' (Figure S1). The descriptors were as follows: pain, rage, disgust, hairs-on-end, headache, flinching, tingling, soothing, pleasurable, annoyance, nausea, visual experiences, distress, body tension, anxiety, discomfort, too loud.

After completing the rating task, the participants took part in a set of other questionnaires. For our main study (N=418), these measures are written-up and analysed separately<sup>14</sup> although we make additional use of some of them here: namely, the autism spectrum quotient AQ,<sup>40</sup> Glasgow Sensory Questionnaire GSQ,<sup>18</sup> and the Multidimensional Assessment of Interoceptive Awareness MAIA,<sup>19</sup>. For our second sample (N=254), the main measures of interest are the previously described SMS and ASMR-15 ASMR-15,<sup>23</sup>. The ASMR-15 consists of fifteen questions requiring an answer on a five-point

Likert scale (Completely untrue for me; Somewhat untrue for me; Neither true nor untrue for me; Somewhat true for me; Completely true for me). They were given the generic instructions: "This survey is looking at how certain stimuli affect you. Some individuals experience intense physical and emotional responses upon hearing particular sounds. These sensations and feelings can be pleasant or unpleasant. Sounds such as whispering, crackling, tapping, or scratching may produce particular experiences described below. Using the scale, please indicate your level of agreement with each statement, upon hearing any of these, or similar sounds:" and then asked to consider it in the context of "When I hear certain sounds, such as whispering, crinkling, tapping . . .". Example items include "It feels like goosebumps on the back of my head" and "The experience is blissful." Items are averaged giving a 1-5 continuous scale.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Machine Learning

An aim of this analysis is to extract, in a data-driven way, a misophonic profile from the 32 sounds (creaking, crunching, etc.) x 17 responses (Rage, Disgust, Soothing, etc.). Initially each sound was analysed separately to consider the possibility that different sounds have a different misophonic profile. The outcome of this stage is a ranked order of the 32 sounds according to their ability to classify misophonics from controls. The dataset consists of, for each sound, a set of 17 features (ratings for each of the dimensions) from all respondents together with category labels (misophonic, non-misophonic). A Random Forest classifier was trained using 10-fold cross validation and one tuned hyper-parameter (the number of features in each tree varying between 2-8). Random Forests have several advantages: they can be used where the number of features exceeds the number of examples without requiring data reduction, they have few hyper-parameters to tune, and the use of ensembles protect against over-fitting. All features were rescaled using the max-min approach. The main dependent variable was accuracy (measured as area-under-curve, AUC) where 0.5 is chance and 1.0 is perfect classification. Specificity (proportion of misophonics correctly classed) and sensitivity (proportion of non-misophonics correctly classed) are noted. Permuted datasets (N=1000) in which group labels are randomly shuffled were created to assess whether the AUC is significantly greater than chance ( $p < .05$ ). When appropriate, FDR (False Discovery Rate), correction for multiple comparisons is also performed<sup>41</sup>. The most important features that drive classification was determined using the R package, VSURF (Variable Selection using Random Forests)<sup>22</sup> which has been shown to perform favourably compared to other solutions<sup>42</sup>. This ranks the N = 17 features by their degree of importance for classification and we report, for interpretation purposes, those features that are not eliminated in the first step.

To analyse the degree to which the ability to classify a person as misophonic is sound-specific, one can apply cross-classification. For example, one can determine whether the 'apple crunching' classifier also be used to identify misophonic people's responses to the 'pen clicking' sound. Here we took the trained classifiers from above and tested them on the response profile from all other sounds taking AUC as a measure of classification accuracy.

The extent to which the putative ‘misophonia profile’ is specific to misophonia or related to other conditions/traits is explored by splicing our participant groups in other ways. This is feasible because our participants completed a variety of other measures reported elsewhere. Specifically, within our control (non-misophonic) group we split the group into high/low according to their scores on the AQ, GSQ and MAIA. For the AQ, we used the published cut-off of  $\geq 23$  for the ‘Broad Autism Phenotype’ and above<sup>17</sup> (yielding groups of  $N = 57$  and 139) and for the other measures a median split was taken in the absence of agreed cut-offs. Within the misophonics, we were interested in the co-morbid presence of hyperacusis. We had previously included a self-report question about hyperacusis (“Hyperacusis: When everyday sounds feel overwhelming, loud, intense, or painful that do not bother other people in the same way”), taken from a scoping review<sup>15</sup>. There were  $N=122$  misophonics agreeing and  $N=99$  disagreeing that this statement applied to them. Another potential measure of hyperacusis, is the Pain factor of the SMS<sup>21</sup> which consists of 4 questions answered on a 5-point Likert scale (never, hardly ever, sometimes, often, always). Here we divided participants according to those answering, on average, above or below the scale midpoint such that the high group was above ‘sometimes’ and the low group was at or below that point (82 and 139 participants respectively). For each of these splits of the dataset, the machine learning analysis procedure described above was re-run although we adopted 5-fold cross-validation to take into account the smaller sample sizes.

### Psychoacoustic Properties of the Sounds

For several analyses, it was important to know more about the psychoacoustic features of our sound stimuli. We applied the mathematical model of Chi, Ru and Shamma<sup>36</sup> on the spectral and temporal features of our sounds to generate a model of the cortical representation of sound filtered/tuned to different spectro-temporal modulations. This was done using the NSL (Neural Systems Laboratory) auditory-cortical Matlab toolbox (<http://nsl.isr.umd.edu/downloads.html>). Specifically, the cortical representation of each sound was represented as four dimensions: sound frequency (128 bins with centre frequencies ranging from 182 to 7143 Hz); spectral modulations (7 bins: 0.125, 0.25, 0.5, 1, 2, 4, 8 cycles per octave), temporal modulations (12 bins: +/- 32, 16, 8, 4, 2 and 1 Hz), and time (5 ms bins). Following the NSL code, the centre frequencies were calculated as follows:  $440 * 2^{[(0:128)-31]/24}$  (this covers around 5.3 octaves with 24 bins per octave, with 440 Hz being a central reference frequency for music). Following Kumar et al.<sup>2</sup> we calculated, for each sound, the mean intensity between 2500-5500 Hz and 1-16 Hz temporal modulations (averaging over time and spectral modulations). These values were correlated against the mean group ratings for the same sounds. We also conducted the same psychoacoustic analysis of the  $N=125$  sounds given to misophonics and controls by Hansen et al.<sup>6</sup> and rated on the single dimension of discomfort. These analyses were not pre-registered and are a post-hoc exploration of the data.

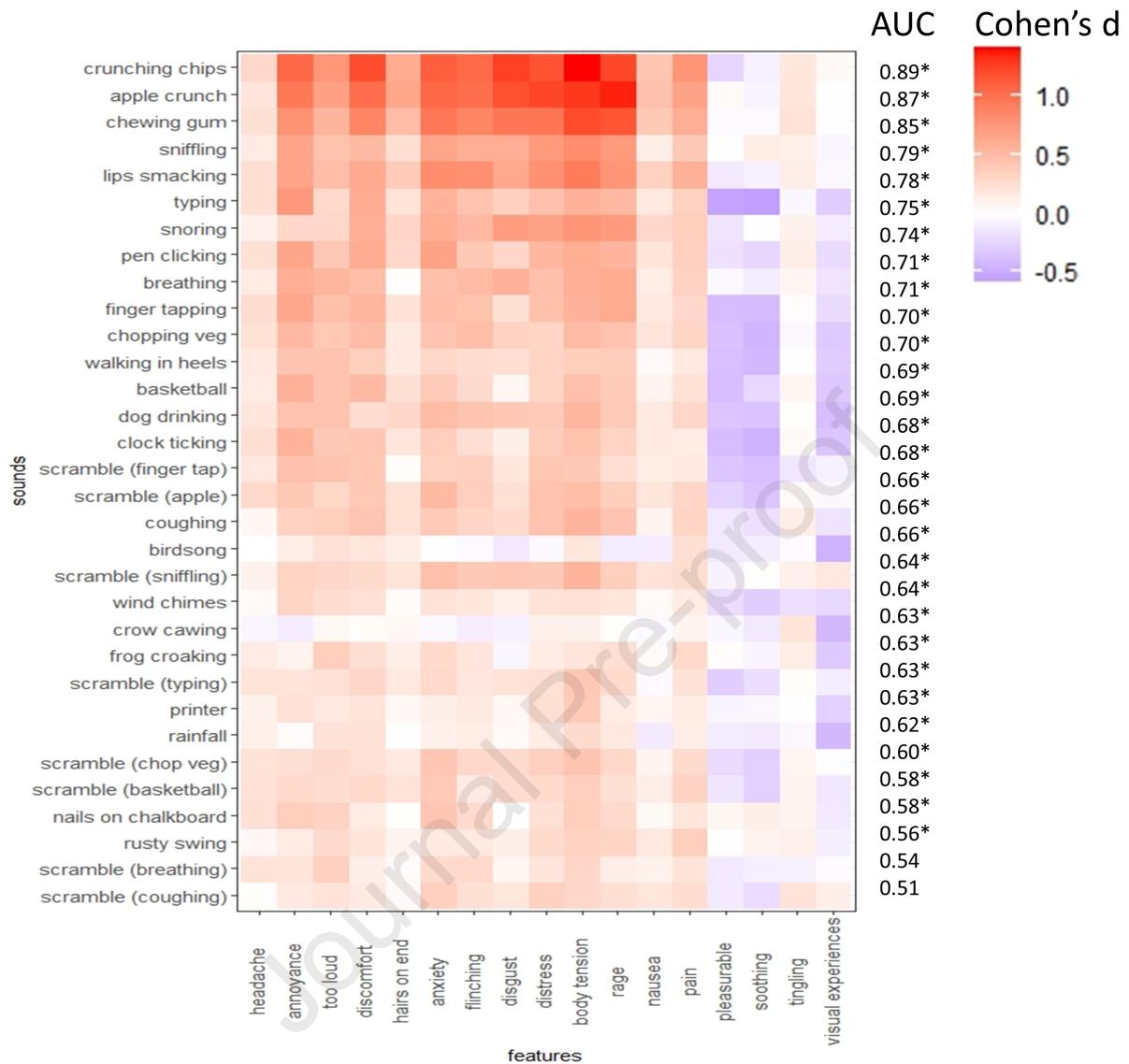
Finally, we also used the derived cortical representations to better understand the effect that our scrambling procedure had on the sounds by comparing the original 24 sounds against their scrambled counterparts. A summary of this analysis is included in the supplemental information (Figures S2 to S5).

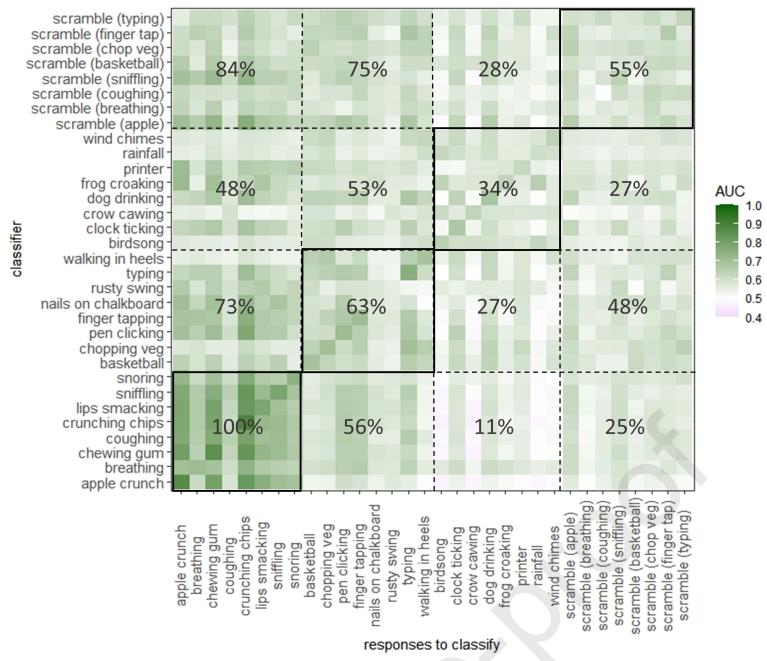
## REFERENCES

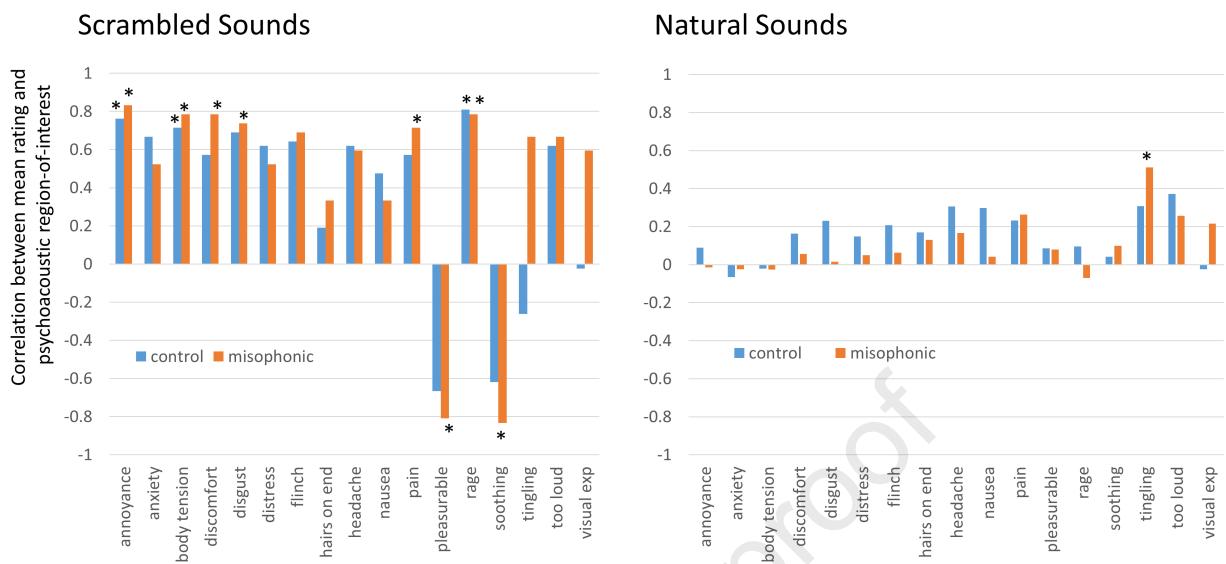
1. Swedo, S.E., Baguley, D.M., Denys, D., Dixon, L.J., Erfanian, M., Fioretti, A., Jastreboff, P.J., Kumar, S., Rosenthal, M.Z., Rouw, R., et al. (2022). Consensus Definition of Misophonia: A Delphi Study. *Frontiers in Neuroscience* 16, 841816. 10.3389/fnins.2022.841816.
2. Kumar, S., Forster, H.M., Bailey, P., and Griffiths, T.D. (2008). Mapping unpleasantness of sounds to their auditory representation. *The Journal of the Acoustical Society of America* 124, 3810-3817.
3. Brout, J.J. (2022). A Brief Commentary on the Consensus Definition of Misophonia. *Frontiers in Neuroscience* 16, 879070. 10.3389/fnins.2022.879070.
4. 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. *Current Biology* 27, 527-533. 10.1016/j.cub.2016.12.048.
5. Schröder, A., van Wingen, G., Eijsker, N., San Giorgi, R., Vulink, N.C., Turbyne, C., and Denys, D. (2019). Misophonia is associated with altered brain activity in the auditory cortex and salience network. *Scientific Reports* 9, 7542. doi.org/10.1038/s41598-019-44084-8.
6. Hansen, H.A., Leber, A.B., and Saygin, Z.M. (2021). What sound sources trigger misophonia? Not just chewing and breathing. *Journal of Clinical Psychology* 77, 2609-2625. 10.1002/jclp.23196.
7. Williams, Z.J., He, J.L., Cascio, C.J., and Woynaroski, T.G. (2021). A review of decreased sound tolerance in autism: Definitions, phenomenology, and potential mechanisms. *Neuroscience and Biobehavioral Reviews* 121, 1-17. 10.1016/j.neubiorev.2020.11.030.
8. Dozier, T.H., and Morrison, K.L. (2017). Phenomenology of Misophonia: Initial Physical and Emotional Responses. *American Journal of Psychology* 130, 431-438. 10.5406/amerjpsyc.130.4.0431.
9. Barratt, E.L., and Davis, N.J. (2015). Autonomous Sensory Meridian Response (ASMR): a flow-like mental state. *Peerj* 3, e851. 10.7717/peerj.851.
10. Niven, E.C., and Scott, S.K. (2021). Careful whispers: when sounds feel like a touch. *Trends in Cognitive Sciences* 25, 645-647. 10.1016/j.tics.2021.05.006.
11. Rouw, R., and Erfanian, M. (2018). A Large-Scale Study of Misophonia. *Journal of Clinical Psychology* 74, 453-479. 10.1002/jclp.22500.
12. McErlean, A.B.J., and Banissy, M.J. (2018). Increased misophonia in self-reported Autonomous Sensory Meridian Response. *Peerj* 6, e5351. 10.7717/peerj.5351.
13. Barratt, E.L., Spence, C., and Davis, N.J. (2017). Sensory determinants of the autonomous sensory meridian response (ASMR): understanding the triggers. *Peerj* 5, e3846. 10.7717/peerj.3846.
14. Andermane, N., Bauer, M., Simner, J., and Ward, J. (2022). A symptom network model of misophonia: From heightened sensory sensitivity to clinical comorbidity. <https://psyarxiv.com/t96wz/>.
15. Fackrell, K., Potgieter, I., Shekhawat, G.S., Baguley, D.M., Sereda, M., and Hoare, D.J. (2017). Clinical Interventions for Hyperacusis in Adults: A Scoping Review to Assess the Current Position and Determine Priorities for Research. *Biomed Research International* 2017, 2723715. 10.1155/2017/2723715.
16. Baguley, D.M., and Hoare, D.J. (2018). Hyperacusis: major research questions. *HNO* 66, 358-363.
17. Wheelwright, S., Auyeung, B., Allison, C., and Baron-Cohen, S. (2010). Defining the broader, medium and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Molecular Autism* 1, 10. 10.1186/2040-2392-1-10.

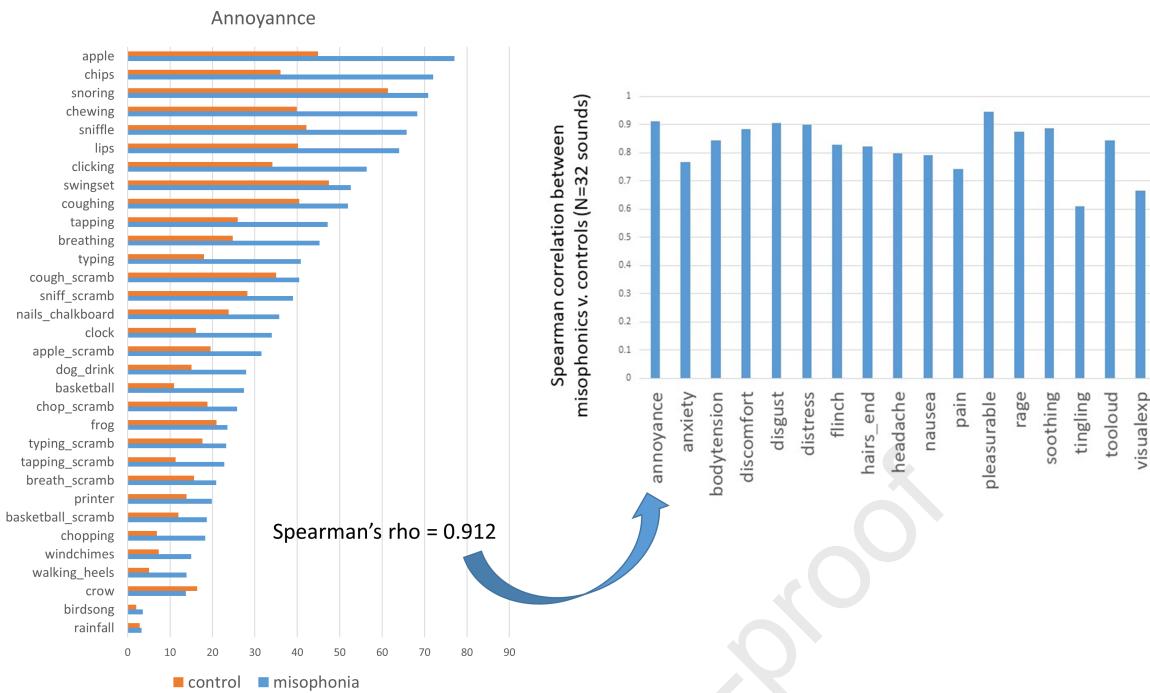
18. Robertson, A.E., and Simmons, D.R. (2013). The Relationship between Sensory Sensitivity and Autistic Traits in the General Population. *Journal of Autism and Developmental Disorders* 43, 775-784. 10.1007/s10803-012-1608-7.
19. Mehling, W.E., Price, C., Daubenmier, J.J., Acree, M., Bartmess, E., and Stewart, A. (2012). The Multidimensional Assessment of Interoceptive Awareness (MAIA). *Plos One* 7, e48230. 10.1371/journal.pone.0048230.
20. 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. *Frontiers in Psychology* 12, 709928. 10.3389/fpsyg.2021.709928.
21. Rinaldi, L.J., Ward, J., and Simner, J. (2022). A Factor Structure within Misophonia: The Sussex Misophonia Scale for researchers and clinicians.
22. Genuer, R., Poggi, J.M., and Tuleau-Malot, C. (2015). VSURF: An R Package for Variable Selection Using Random Forests. *R Journal* 7, 19-33.
23. Roberts, N., Beath, A., and Boag, S. (2019). Autonomous Sensory Meridian Response: Scale Development and Personality Correlates. *Psychology of Consciousness-Theory Research and Practice* 6, 22-39. 10.1037/cns0000168.
24. Gillmeister, H., Succi, A., Romei, V., and Poerio, G.L. (2022). Touching you, touching me: Higher incidence of mirror-touch synesthesia and positive (but not negative) reactions to social touch in Autonomous Sensory Meridian Response. *Consciousness & Cognition* 103. doi: 10.1016/j.concog.2022.103380.
25. Williams, Z.J., Suzman, E., and Woynaroski, T.G. (2021). A Phenotypic Comparison of Loudness and Pain Hyperacusis: Symptoms, Comorbidity, and Associated Features in a Multinational Patient Registry. *American Journal of Audiology* 30, 341-358. 10.1044/2021\_aja-20-00209.
26. Brout, J.J., Edelstein, M., Erfanian, M., Mannino, M., Miller, L.J., Rouw, R., Kumar, S., and Rosenthal, M.Z. (2018). Investigating Misophonia: A Review of the Empirical Literature, Clinical Implications, and a Research Agenda. *Frontiers in Neuroscience* 12, 36. 10.3389/fnins.2018.00036.
27. 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. *Frontiers in Psychology* 13, 890829. 10.3389/fpsyg.2022.890829.
28. Savard, M.A., Sares, A.G., Coffey, E.B.J., and Deroche, M.L.D. (2022). Specificity of Affective Responses in Misophonia Depends on Trigger Identification. *Frontiers in Neuroscience* 16, 879583. 10.3389/fnins.2022.879583.
29. Williams, Z.J., Cascio, C.J., and 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. 10.3389/fpsyg.2022.897901.
30. Aazh, H., Erfanian, M., Danesh, A.A., and Moore, B.C.J. (2022). Audiological and Other Factors Predicting the Presence of Misophonia Symptoms Among a Clinical Population Seeking Help for Tinnitus and/or Hyperacusis. *Frontiers in Neuroscience* 16, 900065. 10.3389/fnins.2022.900065.
31. Kaufman, A.E., Weissman-Fogel, I., Rosenthal, M.Z., Neeman, R.K., and Bar-Shalita, T. (2022). Opening a window into the riddle of misophonia, sensory over-responsiveness, and pain. *Frontiers in Neuroscience* 16, 907585. 10.3389/fnins.2022.907585.
32. Rinaldi, L.J., Simner, J., Koursarou, S., and Ward, J. Autistic traits, emotion regulation, and sensory sensitivities in children and adults with Misophonia. *Journal of Autism and Developmental Disorders*. 10.1007/s10803-022-05623-x.
33. 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. *Journal of Clinical Psychology* 70, 994-1007. 10.1002/jclp.22098.

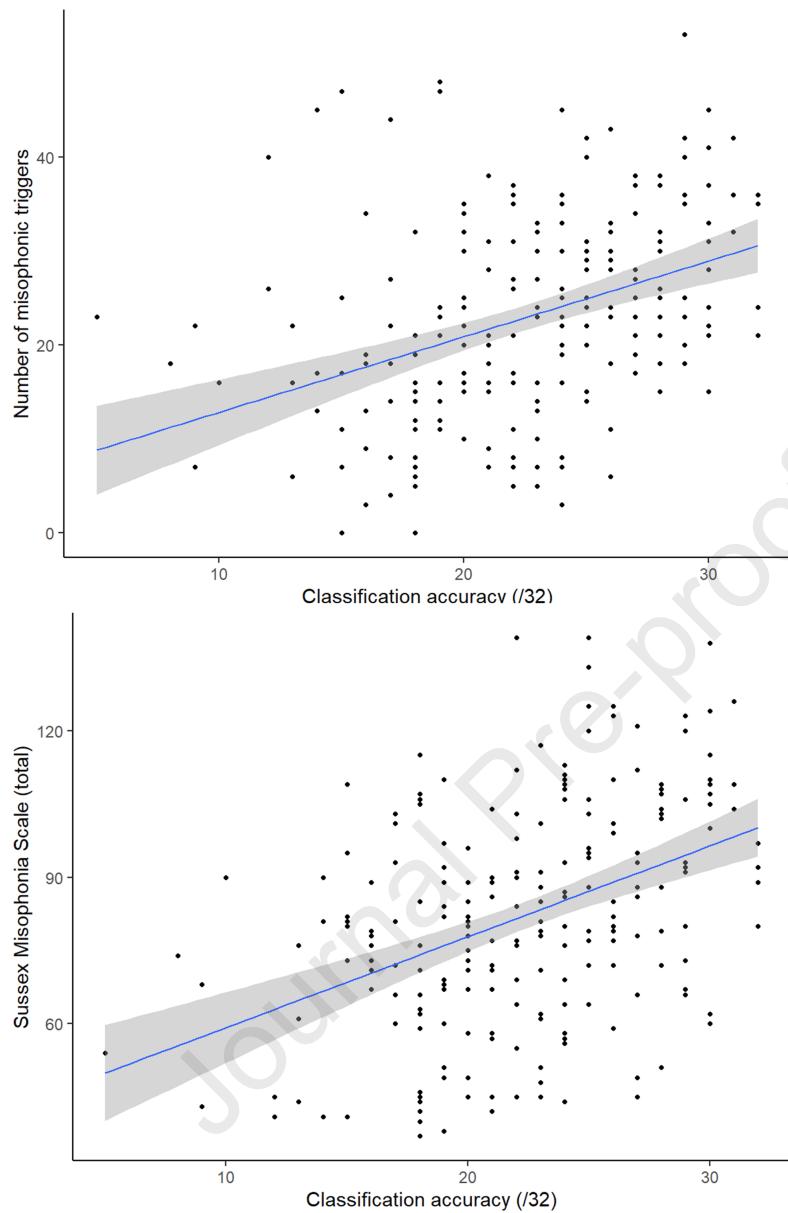
34. Ward, J. (2011). Visual music in arts and minds: Explorations with synesthesia. In Art and the Senses, F. Bacci, and D. Melcher, eds. (Oxford University Press).
35. Ellis, D.P.W. (2010). Time-domain scrambling of audio signals in Matlab. <https://www.ee.columbia.edu/~dpwe/resources/matlab/scramble/#1>.
36. Chi, T., Ru, P.W., and Shamma, S.A. (2005). Multiresolution spectrotemporal analysis of complex sounds. Journal of the Acoustical Society of America 118, 887-906. 10.1121/1.1945807.
37. Jakubovski, E., Müller, A., Kley, H., de Zwaan, M., and K., M.-V. (2022). Prevalence and clinical correlates of misophonia symptoms in the general population of Germany. Frontiers in Psychiatry 13. doi:10.3389/fpsyg.2022.1012424.
38. Edelstein, M., Brang, D., Rouw, R., and Ramachandran, V.S. (2013). Misophonia: physiological investigations and case descriptions. Frontiers in Human Neuroscience 7, 296. 10.3389/fnhum.2013.00296.
39. Boersma, P., and Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.50, retrieved 20 June 2021 from <http://www.praat.org/>.
40. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. Journal of Autism and Developmental Disorders 31, 5-17. 10.1023/a:1005653411471.
41. Benjamini, Y., and Hochberg, Y. (1995). CONTROLLING THE FALSE DISCOVERY RATE - A PRACTICAL AND POWERFUL APPROACH TO MULTIPLE TESTING. Journal of the Royal Statistical Society Series B-Statistical Methodology 57, 289-300.
42. Speiser, J.L., Miller, M.E., Tooze, J., and Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. Expert Systems with Applications 134, 93-101. 10.1016/j.eswa.2019.05.028.

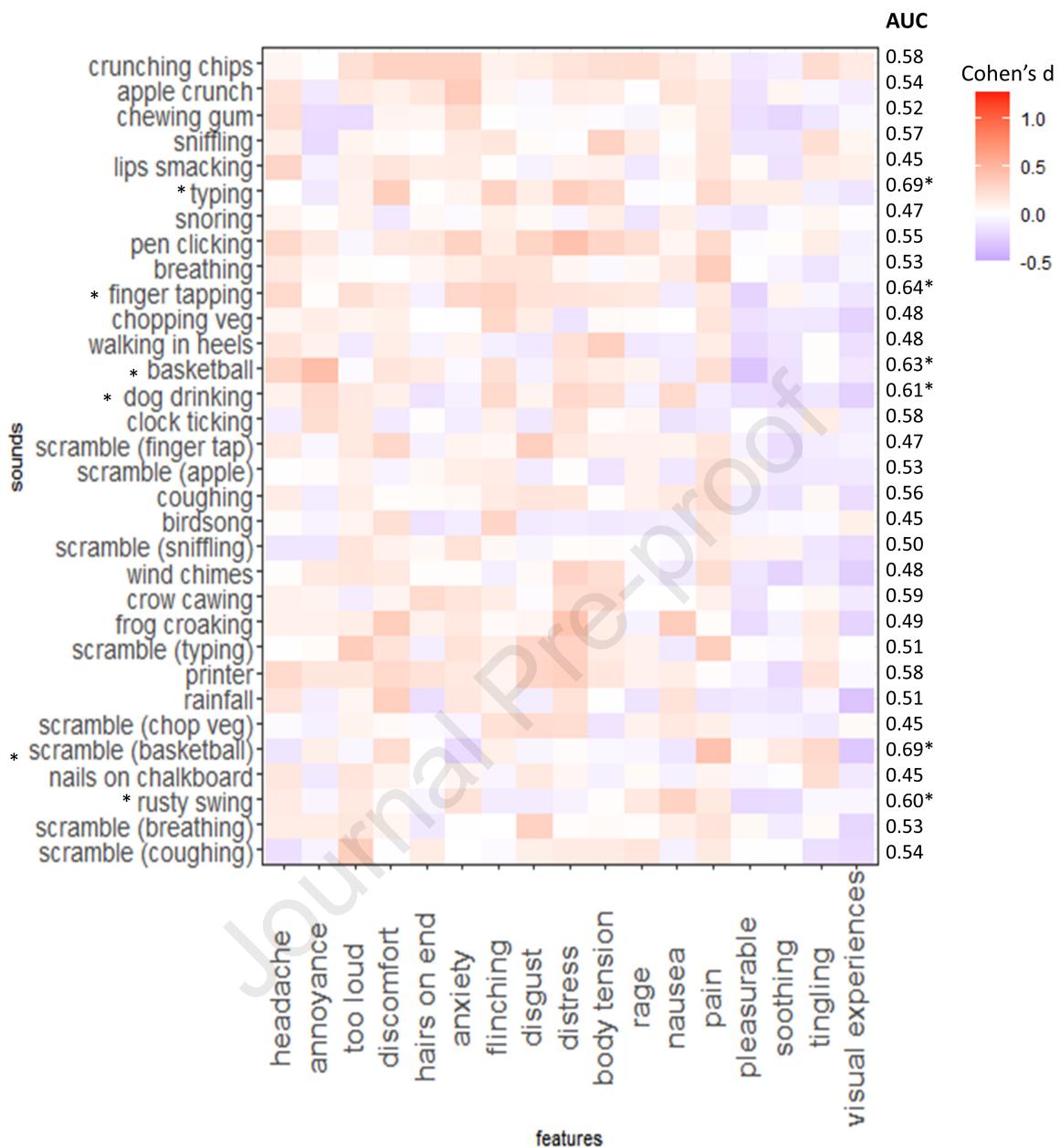




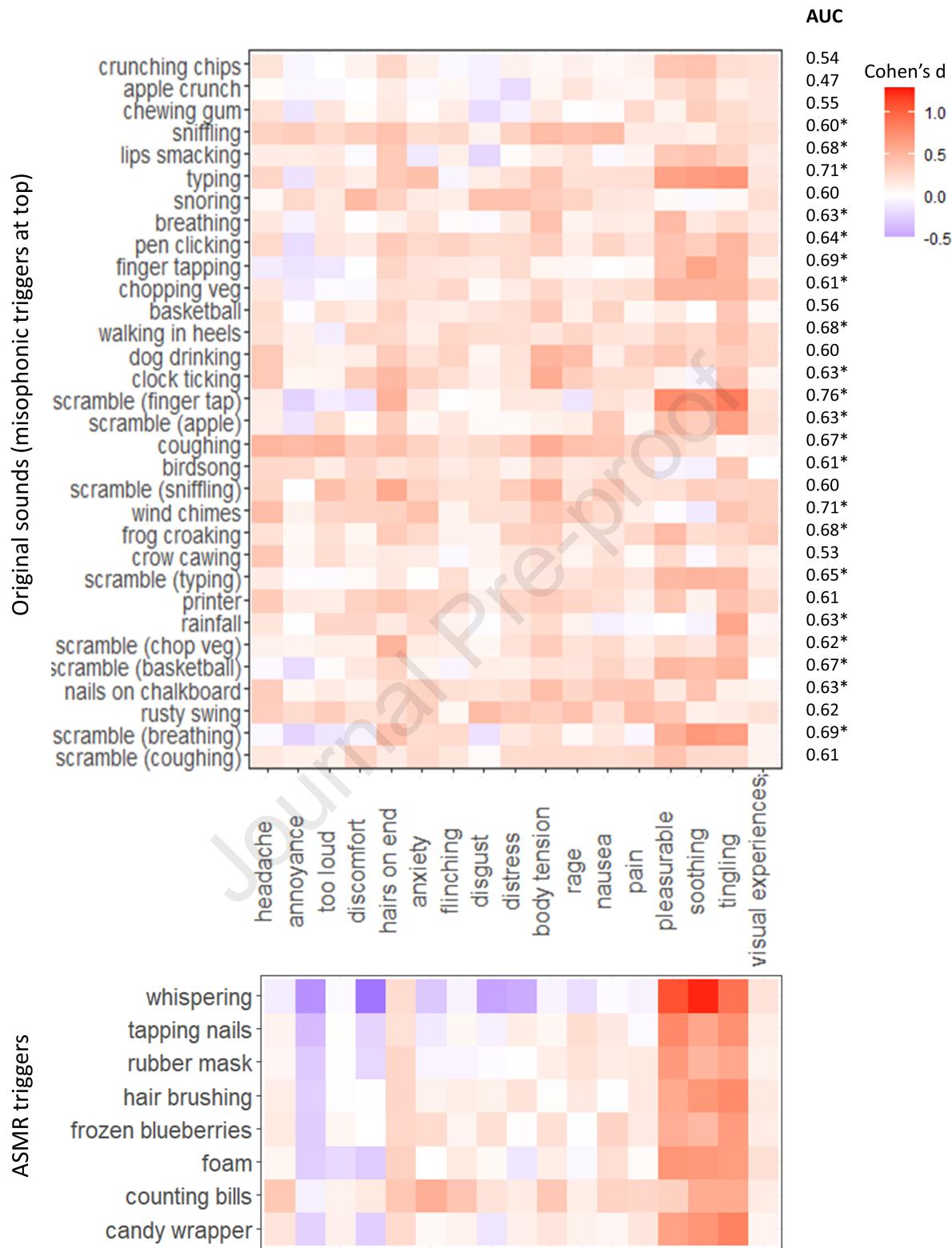












**Highlights**

- Misophonia can be identified by applying machine learning to a sound-response profile
- Misophonia is linked to broad sound intolerances beyond specific triggers
- Sound intolerances linked to autism and hyperacusis are different to misophonia
- Our findings do not support a link between misophonia and ASMR