Phenomenology of Misophonia: Initial Physical and Emotional Responses

Article in The American Journal of Psychology · January 2017
DOI: 10.5406/amerjpsyc.130.4.0431

1 author:

Thomas Dozier
Misophonia Treatment Institute
8 PUBLICATIONS 21 CITATIONS

Some of the authors of this publication are also working on these related projects:

- Misophonia Initial Physical Reflex Measurement View project
- Misophonia Research View project

All content following this page was uploaded by Thomas Dozier on 11 November 2017.

The user has requested enhancement of the downloaded file.
Misophonia is typically characterized as an extreme emotional response to auditory and visual stimuli. In several case studies, physical responses have also been reported. This study sought to determine whether adults with misophonia experience physical responses in addition to emotional responses in the presence of triggering stimuli. Twenty-seven adults with misophonia were interviewed via teleconferencing. Participants self-reported the presence of physical and emotional responses to triggers (i.e., two auditory and one visual). All participants reported physical responses to at least 1 of their triggers. There was great variation in the region of the physical responses across participants. Approximately half of the sample reported region consistency across triggers. Likewise, all participants reported emotional responses to at least 1 of their triggers. These results suggest including an immediate physical response as part of the conceptualization of misophonia. They also support classical conditioning of a physical response as a possible contributing mechanism for the etiology of misophonia.

KEYWORDS: misophonia, classical conditioning, reflex, selective sound sensitivity syndrome, neurodevelopmental disorders, Pavlovian conditioning, sound sensitivity
Several studies have reported physiological arousal in response to trigger stimuli, including increased skin conductance, heart rate, blood pressure, and a feeling of pressure in the body, as an accompanying stress response to misophonic triggers (Claiborn et al., 2017; Edelstein et al., 2013; Jastreboff & Jastreboff, 2014; Schröder et al., 2013). Edelstein et al. (2013) empirically validated the autonomic arousal by measuring the skin conductance response (SCR) to prolonged exposure to misophonic auditory stimuli. SCR began rising 2 s after onset of the trigger stimulus and continued to rise for the duration of the stimulus. The stress response is illustrated in Figure 1a, which displays the prevalent conception of misophonia.

Several case studies and one survey have reported a physical response, typically a skeletal muscle contraction (e.g., flinch, various muscle contractions), to trigger stimuli (Claiborn et al., 2017; Dozier, 2015a, 2015b, 2015c; Pearson, 2015). Dozier (2015b, 2015c, 2015d) proposed an alternative model of the misophonic response that includes this physical response elicited by the trigger stimulus (Figure 1b). In the proposed model, the misophonic emotional response is elicited primarily by the sensation of the initial physical response. The dashed line in Figure 1b shows a secondary path whereby the trigger stimulus may directly contribute to the strength of the emotional response. The proposed response paths are similar to the James–Lange and Cannon–Bard theories of emotion, which discuss the order and processing of physiological and emotional responses to stimuli (Cannon, 1987). Though similar, the model put forth by Dozier (2015b, 2015c, 2015d) includes a sometimes visible muscle flinch or ticlike movement beyond the emotional and physiological responses in those models. Dozier’s model proposes that misophonia is an aversive physical and emotional reflex disorder. Moreover, Dozier (2015b) posits that misophonia includes a classically conditioned aversive physical response to selective stimuli, which does not extinguish with in vivo exposure.

Aversive stimuli have been shown to evoke anger in humans (Berkowitz, 1983; Berkowitz, Cochran, & Embree, 1981). Neuroimaging studies have shown increased activity in the limbic system of humans in response to aversive gustatory stimuli (Zald, Lee, Fluegel, & Pardo, 1998), aversive odorants (Zald & Pardo, 1997), and aversive auditory stimuli (Zald & Pardo, 2002). The physical response to misophonic trigger stimuli has been reported as severe by some people and very mild by others (Dozier, 2015b). Even in cases where the physical response is weak, it is intrusive and may therefore be aversive.

The existence of an initial physical response is an important consideration in elucidating misophonia, but the emotional response to real-life trigger stimuli may mask the perception of that response (Dozier, 2015c, 2015d). This study sought to determine whether physical sensations could be identified when misophonic people were presented with weak trigger stimuli. Providing evidence of an initial physical response in a controlled study may lead to increased awareness and recognition of this previously reported characteristic of the misophonic response. It may also identify a widespread example where a physical response contributes to extreme emotional responses.

**FIGURE 1.** (a) Prevalent conception of misophonia emotional and physical response to trigger stimuli. The trigger stimulus elicits the misophonic extreme emotions and the accompanying stress response. (b) Alternative conception of misophonia response chain, including the initial physical response. The trigger stimulus elicits a physical response, and the sensation of the physical response elicits the extreme emotions and accompanying stress response.
EXPERIMENT

METHOD

Participants
Invitations were sent to 194 people (31 men and 163 women) randomly selected from those who, as part of an online survey of adults with misophonia symptoms, indicated a willingness to participate in further study phases. The high rate of female participation in the survey is consistent with other self-selection misophonia studies (Cash, 2015; Wu et al., 2014), even though those studies found no significant gender difference for prevalence. Eligibility required that participants be 18 years of age or older and report at least two misophonic trigger stimuli. Of the 194 contacted, 27 people returned the informed consent form and completed the study interview (7 men and 20 women, mean age = 38.3 years; range = 18–63 years). The responses of one person were excluded from analysis because responses intermixed thoughts, emotions, and physical sensations for both the physical and emotional questions. Based on the Amsterdam Misophonia Scale (range 0–24; Schröder et al., 2013), which was administered as part of a previous study, the misophonia severity ratings of participants ranged from subclinical to extreme, with the mean between moderate and severe ($M = 14.38, SD = 3.81$, range = 4–21).

Procedure
This study was approved by the Sterling Institutional Review Board. After being invited to participate in the interview, participants returned the completed informed consent, a form that designated the types of triggers that would be used in the interview and, if necessary, any recording of those triggers via e-mail. The trigger form requested identification of one visual and two auditory triggers. Each trigger was designated as either “trigger made by investigator” or “recording will be provided.” At that time, a one-time appointment was scheduled to meet via VSee (a Health Insurance Portability and Accountability Act–compliant telehealth videoconferencing platform). Interviews were conducted by both authors.

Before the interview, the interviewer verified that recorded trigger files played properly and prepared the materials to produce triggers (e.g., chips, gum, video of kissing). The interviews lasted approximately 20–30 min. At the start of the interview, the informed consent was reviewed for participant clarification. Once assured it was fully understood, the interviewer read the following script:

The purpose of this meeting is to help you identify the physical sensations that you have when you are triggered. In real life, the emotions are so overpowering that almost no one can accurately identify the physical sensations that occur immediately after a trigger. The research on reflexes indicates that the smaller a trigger stimulus, the smaller the response. I will attempt to make the trigger so tiny that it does not cause any reflex response. Then I will slowly increase the trigger until you start to be triggered. At that point, I will ask you to report the strength of the trigger (on a scale of 0–10), any physical sensation you have, and any emotions you experience.

Each trigger was tested several times. For each auditory trigger test, participants were asked to close their eyes and relax their muscles. The interviewer produced a very low-level trigger (i.e., low volume and short duration), and the participant was asked whether it triggered them. The trigger was slowly increased until the participant was triggered. The participant then reported the strength of the trigger, physical sensations, and emotions experienced. The interviewer prompted for clarification of the responses when needed. For subsequent tests of the same trigger, the interviewer attempted to adjust the trigger to have a “trigger strength” of 2 to 4, as reported by the participant. This procedure was repeated several times until a consistent physical response was reported (defined as two matching consecutive responses) or a maximum of 10 tests were completed. The procedure for visual triggers was the same as for auditory triggers, except participants were asked to relax their muscles with their eyes open and close their eyes as soon as they were triggered. The researcher produced the trigger, attempting to start with a nontriggering stimulus and slowly increase the trigger strength until participants closed their eyes. The strength of the visual trigger was adjusted by the size of the image on the screen and topology. For example, if the trigger was open mouth chewing, the interviewer started chewing with a closed mouth and progressively increased the separation of the lips until the participant closed his or her eyes. With this method, the trigger strength rating was maintained.
in the same range for visual and auditory triggers. The interviewer’s microphone was muted during the test for a visual trigger to ensure that the misophonic response was elicited solely by the visual image. The two auditory and one visual triggers were tested using this format, with the exception of two participants who had only auditory triggers. Participant responses were recorded on a data collection form by the researcher and later entered into a data file.

Interviewer Interrater Reliability
To ensure interviewer interrater reliability, 6 of the 27 interviews (22%) were observed by the other author. The observer was blocked visually and audibly from both the participant and interviewer after making the participant aware of his or her presence. The observer completed the same data collection form and entered the data into a separate data file for later comparisons. Interrater agreement was calculated for the physical and emotional response data. The percentage agreement was calculated as the number of responses that completely matched, divided by total responses. Interrater agreement was 91.1%, which suggests high agreement between the observer and interviewer.

Data Analysis Techniques
The physical response was determined according to three rules. First, if two consecutive responses were identical, those responses were reported (57 of 76 tests, 75.0%). Second, if a set of responses did not meet the first rule but were reported in three consecutive tests, and each response was reported at least twice, then all the responses were reported (9 of 76 tests, 11.8%). Third, if the first and second rules were not met, responses that occurred two or more times were reported, and any further responses were reported as various (10 of 76 tests, 13.2%). All emotional responses reported were included in the data summary. The second author reviewed 19.2% (randomly selected) of the total participant responses to ensure accuracy of the data summary. The data summary interrater agreement was 100%.

RESULTS
Physical Responses
Each participant was tested with two different auditory triggers and one visual trigger, except for the two participants who did not have any visual misophonic triggers. The physical and emotional responses for all triggers tested are shown in the supplemental table. The physical responses for each person are summarized in Table 1. Responses are grouped into major categories. For example, responses of upper arms, forearms, hands, and fingers are grouped into arms and hands. A response of trapezius (e.g., participant reporting shoulders at the neck or pointing to this location) is included in the shoulders and neck groups.

Three participants (11.5%) had one or two triggers wherein they consistently reported no physical response. One participant (3.8%) reported a physical response but intermittently reported no physical response from two triggers. The responses to many
of the triggers included more than one muscle group; therefore, the sum of the percentages in Table 1 is greater than 100%. Table 2 shows the number of participants with same, similar, different, and no physical responses. All participants (100%) reported a physical response to at least one of their tested triggers. Table 3 provides examples of the consistency of responses to different trigger stimuli.

It was difficult for some participants to determine the location of the response. For example, one participant reported feeling panic and fear in the first instance of the trigger and a physical sensation in the gut and shoulders in the second instance. After several more exposures to weak triggers, the participant consistently reported a slight tensing of the abdomen as the physical response to triggers. Therefore, in many cases the exact location of the physical response should not be considered reliable (e.g., whether the response was in the neck versus the neck, shoulders, and chest).

**Emotional Responses**

Emotional responses were recorded for each trigger test. Actual responses from each trigger tested are listed in the supplemental table. Four participants (15.4%) had one or two triggers wherein they consistently reported no emotional response. The only emotion that one of these participants reported in all individual trigger tests was that he “wanted them to go away.” Another four participants (15.4%) reported emotional responses but intermittently reported no emotional response for one or more triggers. A summary of the emotional responses to triggers is shown in Table 4, as a percentage of participants and of triggers tested. As with the physical responses, the sum of the percentages is greater than 100% because multiple emotions were often reported for each trig-

---

**TABLE 2. Within-Participant Similarities and Differences of Responses to Triggers**

<table>
<thead>
<tr>
<th>Responses to different triggers</th>
<th>Number of participants (n = 26)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same (all)</td>
<td>6</td>
<td>23.1%</td>
</tr>
<tr>
<td>Same or similar (all)</td>
<td>15</td>
<td>57.7%</td>
</tr>
<tr>
<td>Same (2 or more)</td>
<td>13</td>
<td>50.0%</td>
</tr>
<tr>
<td>Same or similar (2 or more)</td>
<td>19</td>
<td>73.1%</td>
</tr>
<tr>
<td>Different (1 or more)</td>
<td>11</td>
<td>42.3%</td>
</tr>
<tr>
<td>No response (1 or more)</td>
<td>4</td>
<td>15.4%</td>
</tr>
<tr>
<td>Physical response (1 or more)</td>
<td>26</td>
<td>100%</td>
</tr>
</tbody>
</table>

**TABLE 3. Examples of Physical Responses Consistency to Triggers**

<table>
<thead>
<tr>
<th>Responses to different triggers</th>
<th>Example of physical response of participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory trigger 1</td>
<td>Hand clench, jaw</td>
</tr>
<tr>
<td>Auditory trigger 2</td>
<td>Hand clench, jaw</td>
</tr>
<tr>
<td>Visual trigger</td>
<td>Hand clench, jaw</td>
</tr>
<tr>
<td>Same (all)</td>
<td>Neck, shoulders, upper back, upper arms</td>
</tr>
<tr>
<td>Same or similar (all)</td>
<td>Shoulder, upper back</td>
</tr>
<tr>
<td>Same (2 or more)</td>
<td>Shoulders, chest</td>
</tr>
<tr>
<td>Same or similar (2 or more)</td>
<td>Varied</td>
</tr>
<tr>
<td>Different (1 or more)</td>
<td>Shoulders, upper arms, chest</td>
</tr>
<tr>
<td>No response (1 or more)</td>
<td>Shoulders, upper body, abdomen</td>
</tr>
<tr>
<td>Physical response (1 or more)</td>
<td>Whole arms</td>
</tr>
</tbody>
</table>

**TABLE 4. Summary of Emotional Responses to Trigger Stimuli**

<table>
<thead>
<tr>
<th>Emotional response</th>
<th>Number of participants (n = 26)</th>
<th>% of Participants</th>
<th>Number of trigger tests (n = 76)</th>
<th>% of Trigger tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>24</td>
<td>92.3%</td>
<td>57</td>
<td>75.0%</td>
</tr>
<tr>
<td>Anxiety</td>
<td>24</td>
<td>92.3%</td>
<td>32</td>
<td>42.1%</td>
</tr>
<tr>
<td>Desire for escape</td>
<td>14</td>
<td>53.8%</td>
<td>23</td>
<td>30.3%</td>
</tr>
<tr>
<td>Disgust</td>
<td>12</td>
<td>46.2%</td>
<td>17</td>
<td>22.4%</td>
</tr>
<tr>
<td>Fear</td>
<td>6</td>
<td>23.1%</td>
<td>9</td>
<td>11.8%</td>
</tr>
<tr>
<td>Sadness</td>
<td>4</td>
<td>15.4%</td>
<td>5</td>
<td>6.6%</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>34.6%</td>
<td>9</td>
<td>11.8%</td>
</tr>
<tr>
<td>None</td>
<td>8</td>
<td>30.8%</td>
<td>15</td>
<td>19.7%</td>
</tr>
</tbody>
</table>
The categories of emotion are aggregations of similar emotions. For example, aggravation, irritation, annoyance, frustration, anger, and rage are combined into the anger category. All participants (100%) reported an emotional response to at least one of their tested triggers.

DISCUSSION

Although other studies have reported general physiological arousal in response to strong or prolonged trigger stimuli including muscle tension (Edelstein et al., 2013; Schröder et al., 2013; Wu et al., 2014), this is the first research study to report a physical response to trigger stimuli that often included specific skeletal muscle contraction and movement. Edelstein et al. (2013) concluded that their misophonic participants “reported physical symptoms synonymous with autonomic arousal” (p. 8). In the current study, participants reported their perceived physical response after weak auditory stimuli. Weak stimuli and time between tests were intended to prevent general physiological arousal, so that the participants could be aware of their physical sensations in response to trigger stimuli. We posit that the reported physical sensation was an immediate elicited response from the stimulus and not autonomic arousal, as reported in other studies. Although there were some trigger tests in this study wherein the participant reported experiencing no physical response, most participants reported a physical response to each trigger stimulus, and every participant reported a physical response to at least one of their tested triggers.

The physical responses were highly varied, with reports of skeletal muscle responses predominating. Some participants reported a response in a single area, such as hands or abdomen, and many reported multiple muscles (e.g., hands and shoulders). The non–skeletal muscle responses included a feeling of warmth, stomach contraction, nausea, clitoral sensation, and contraction of the vagina. More than half of participants reported the same or similar physical response for all triggers tested. Most reported the same or similar physical response for two or more triggers, and 42% of participants had distinctly different physical responses to at least one of their triggers. The most important finding is that the misophonic response to triggers includes a physical response that may be unique to each participant. Thus, the recommendation would not be to categorize misophonia by specific physical responses but rather incorporate the general presence of these automatic physical responses into the understanding of this phenomenon.

Anger (and lesser precursors), anxiety, disgust, and avoidance were the most common emotions reported by participants in response to a single weak instance of a trigger, with anger being the dominant emotion. This was consistent with emotional reports in other studies in response to typical trigger stimuli (Edelstein et al., 2013; Wu et al., 2014). However, the reported fear response was novel because it is generally not reported in misophonia studies (Edelstein et al., 2013; Schröder et al., 2013; Wu et al., 2014).

Weak trigger stimuli were used in this study to limit the emotional responses and allow perception of the physical response. The severity of the responses reported by the participant appeared to be directly related to the volume and duration of the trigger stimulus, and these parameters were varied based on the participant rating of severity. The mild emotional responses (e.g., irritation, frustration, mild anger, mild disgust) were probably a function of the weak trigger stimuli. It may be more appropriate to say that the weak trigger stimuli limited the physical and emotional responses, because the strength of a conditioned physical response is related to conditioned stimulus intensity (Kessen, 1953). It is posited that strong trigger stimuli, especially those in real-life situations, would produce stronger physical and emotional responses. About a third of participants (30.8%) had instances where they did not experience an emotional response, and half as many (15.4%) had instances where they did not experience a physical response. The trigger stimuli were intentionally weak. Each participant had several instances where the auditory stimulus did not elicit a physical or emotional response (i.e., participant did not have a misophonic response to the stimulus). It seems plausible that a physical and an emotional response will almost always be elicited by stronger, real-life trigger stimuli. The occasional lack of emotional response should not alter the conceptualization of misophonia as a condition in which innocuous stimuli elicit or evoke extreme emotional responses.

The variation in physical responses of participants supports an individual learning history to
account for the variety of responses. Classical conditioning is a possible mechanism as part of the etiology of misophonia, in which each person develops unique physical responses to repeating stimuli in their environment. Classical conditioning is typically conceptualized as an association between a conditioned stimulus (CS) and an unconditioned stimulus (US), and when the intermittent temporal relationship between the CS and US stimuli is eliminated, the response extinguishes. However, with misophonia there is no identifiable US for conditioning the misophonic stimulus response, and the response generally does not extinguish (Dozier, 2015b). Donahoe and Vegas (2004) reported that conditioning occurred based on the temporal pairing of a neutral stimulus (NS) and an unconditioned response. Dozier (2015b) provides case study examples of people whose conditioning may have occurred with the pairing of an NS (e.g., chewing sound) and an unconscious behavior. Schroder et al. (2013) hypothesized that recurrent conditioning may occur when a child is exposed to repetitive annoying events related to eating sounds. Such events could pair an NS (chewing sound) and a behavioral response (e.g., fist clench, shoulder shrug) unique to the child. This pairing of NS and response would meet the criteria for conditioning reported by Donahoe and Vegas (2004). Further basic research on stimulus–response conditioning may provide valuable insight and inform future misophonia research. Such basic research may also provide insight into other potential physical–emotional interactions such as hating a portable chemotherapy infusion pump (J. Theobald, personal communication, December 16, 2015; http://newccboard.colonclub.com/viewtopic.php?t=11244) or an aversive muscle response to phone ringtones that have been paired with stressful problems (Dozier, 2015b).

Although the nature of this study is insufficient to determine the relationship of the physical and emotional responses, the presence of physical responses suggests a different definition of misophonia than the prevalent conception of misophonia. Thus, a definition that includes the presence of physical responses is recommended based on the findings from this study: Misophonia is a condition in which a person experiences an immediate physical response and an immediate negative emotional response to auditory, visual, or other modality trigger stimuli.

**Limitations**

This study has inherent limitations due to its reliance on self-report of participants. The stimuli were low intensity and elicited only mildly aversive responses. The stimuli for each person were different based on their individual misophonic triggers. Many of the participants had difficulty identifying a consistent physical response, especially at the start of their interviews, and so the repetitions were different for each participant. Continued practice or exposure to the stimuli could have led to differing responses. Additionally, there is a lack of accuracy in specifying the location of the physical response. Although there may be a lack of precision regarding the location of the physical response, this does not diminish the fact that participants experienced a physical response to exposure of a trigger stimulus.

The method of this study does not provide concrete information about timing of the physical or emotional responses in relation to the trigger stimulus or each other. Additional testing is needed to provide empirical data on the temporal relationship of the stimulus, physical responses, and emotional responses. The relationship between the stimulus and physical response might be measured with electromyography, and the temporal relationship between the stimulus and emotional response might be measured with brain imaging technology. This study supports the existence of a distinct physical response but does not provide information on the priority of the physical response as hypothesized in Figure 1b. The effects of self-selection to participate in this study, the high percentage of female participants, and conducting the study via webcam are unknown.

**Conclusion**

This study provides insight into the composition of misophonia, indicating that misophonia consists of both involuntary physical and emotional responses to trigger stimuli. It also documents that auditory and visual trigger stimuli produce similar physical responses, demonstrating that misophonia may be a general sensory condition rather than an auditory condition. The existence of a physical response has important implications for understanding the origin, maintenance, and treatment of misophonia. The results from this study may suggest conceptualizing misophonia as an aversive physical and emotional
reflex disorder in order to highlight the characteristics of misophonia. Furthermore, this study illustrates a potentially fundamental psychological process, in which a physical response may be the maintaining element of what superficially appears to be a condition in which typically innocuous stimuli elicit extreme negative emotional responses.

NOTES
Thomas Dozier is a misophonia treatment provider and the owner of several misophonia apps.

Address correspondence about this article to Thomas H. Dozier, Misophonia Treatment Institute, 5801 Arlene Way, Livermore, CA 94550 (e-mail: tom@misophoniatreatment.com).

REFERENCES


