

Running head: Misophonia: A Conditioned Aversive Reflex Disorder

Cite: Dozier, T. H. (in press). Etiology, composition, development and maintenance of misophonia: A conditioned aversive reflex disorder. *Psychological Thought*.

Etiology, composition, development and maintenance of misophonia: A conditioned aversive reflex disorder

Thomas H Dozier, United States

tom@misophoniaticreatment.com

Abstract

Misophonia is a recently identified condition in which an individual has an acute reaction of hatred or disgust to a specific commonly occurring sound. We propose that misophonia is a form of conditioned behavior that develops as a physical reflex through classical conditioning. Although misophonia is generally considered to be a one-step reaction, in which the sound elicits rage or disgust, as well as typical autonomic responses associated with these emotions, we propose that misophonia is a two-step reaction, in which the sound elicits an aversive conditioned physical reflex, and the aversive conditioned physical reflex elicits hatred or disgust. We also propose that the emotional response to trigger stimuli creates a classical conditioning paradigm that maintains or strengthens the misophonic physical reflex. Finally, we propose that new misophonic trigger stimuli are developed through the pairing of a neutral stimulus with a misophonic trigger stimulus. We suggest that a better name for misophonia is

Conditioned Aversive Reflex Disorder (CARD) since it focuses attention on the reflexive nature of this condition and incorporates multiple stimuli modalities. A counterconditioning treatment for misophonia is presented with brief case descriptions which demonstrate the conditioned reflex nature of this disorder.

Keywords

misophonia, reflex, conditioning, aversive sounds, conditioned response, counterconditioning, etiology

1. Introduction

Misophonia is characterized by extreme and irrational reactions of hate, anger, rage or disgust which are elicited by soft, commonly occurring sounds such as chewing or sniffing. The response often begins as irritation or disgust, but immediately escalates to extreme emotions. The eliciting stimulus, referred to as a trigger, can be any typically occurring sound, though oral and nasal sounds are most prevalent. The response to triggers is perceived by the individual as involuntary, and individuals commonly report feeling a loss of self-control (Schröder, Vulink, & Denys, 2013). Misophonia is a discrete and independent condition, and does not meet the criteria of any DSM-IV or DSM-5 conditions (Schröder et al., 2013).

This condition was first identified in 1997 by audiologist Marsha Johnson, who labeled it Selective Sound Sensitivity Syndrome or 4S (Bernstein, Angell, & Dehle, 2013). The term misophonia, which literally means dislike or hatred of sound, was proposed by Jastreboff and Jastreboff (2002). 4S more accurately describes the condition since each person has a unique set of specific sounds to which they are highly

sensitive. Although hatred is a common emotion with this disorder, it is not directed at the sound but toward the individual making the sound, and often includes ideation of harm (Schröder et al., 2013). While individuals report anticipatory anxiety for trigger situations, fear of the actual sound is rarely reported (Edelstein, Brand, Rouw, & Ramachandran, 2013; Wu, Lewin, Murphy, & Storch, 2014).

Misophonia typically begins in the preteen years (Edelstein et al., 2013; Wu et al., 2014) but onset has been reported at ages ranging from as young as 2 years to middle age (Edelstein et al., 2013). Initially, trigger stimuli are usually associated with specific individuals, but triggers commonly develop for other people making the sound, variations of the sound, and in other settings (Edelstein et al., 2013). Trigger stimuli include all types of eating sounds (e.g., chewing, lip smacking, crunching, sipping, slurping), nasal sounds (e.g., sniffing, breathing, snoring, nose whistling), and many other sounds (e.g., typing, tapping, consonant sounds, clock ticking, pipes knocking, refrigerator humming, dog barking, sound through walls, footsteps) (Edelstein et al., 2013; Schröder et al., 2013; Wu et al., 2014). Triggers also develop to visual stimuli associated with the sounds (e.g., jaw movement) and repetitive movements not associated with any auditory trigger (e.g., leg jiggling, hand movements) (Wu et al., 2014).

Only one study reported the prevalence of misophonia in a nonclinical sample. That study included 483 undergraduate students (16% male) and found that nearly 20% had clinically significant misophonia symptoms (Wu et al., 2014).

The hallmark of misophonia is an extreme emotional response to the trigger stimulus. One study also reported physiological arousal as being part of the misophonic response:

“On top of the strong psychological effects, misophonics also report experiencing strong physical effects in response to trigger sounds. The most commonly reported physical effects were pressure in the chest, arms, head, or entire body as well as clenched, tightened, and tense muscles. Some misophonics reported an increase in blood pressure, heart rate or body temperature, sweaty palms, physical pain, and even difficulty breathing in response to trigger sounds (see Table 1). The aforementioned aversive responses evoked by trigger sounds are characteristic of a typical, autonomic nervous system response”. (Edelstein et al., 2013, p. 3)

Edelstein et al. also measured skin conductance to assess the autonomic arousal of misophonic participants. They reported that skin conductance began increasing 2 seconds after stimulus onset and continued increasing for the duration of that stimulus. This corroborated the verbal report of participants regarding visceral autonomic responses to trigger stimuli. Jastreboff and Jastreboff (2013; 2014), who have treated individuals with misophonia for over a decade, proposed that misophonia follows the principles of conditioned reflexes. Figure 1 illustrates this view of misophonia, showing that when a trigger stimulus is perceived by the individual, it elicits a response of extreme negative emotions and fight-or-flight responses. Thus far, there has been no research on the etiology of misophonia or specifics of the respondent behavior, but if misophonia is a result of classical conditioning, then, as shown in Figure 1, it is primarily a conditioned emotional response with accompanying physiological arousal (fight-or-flight) to the trigger stimuli. Based on work with patients, we propose alternate hypotheses.

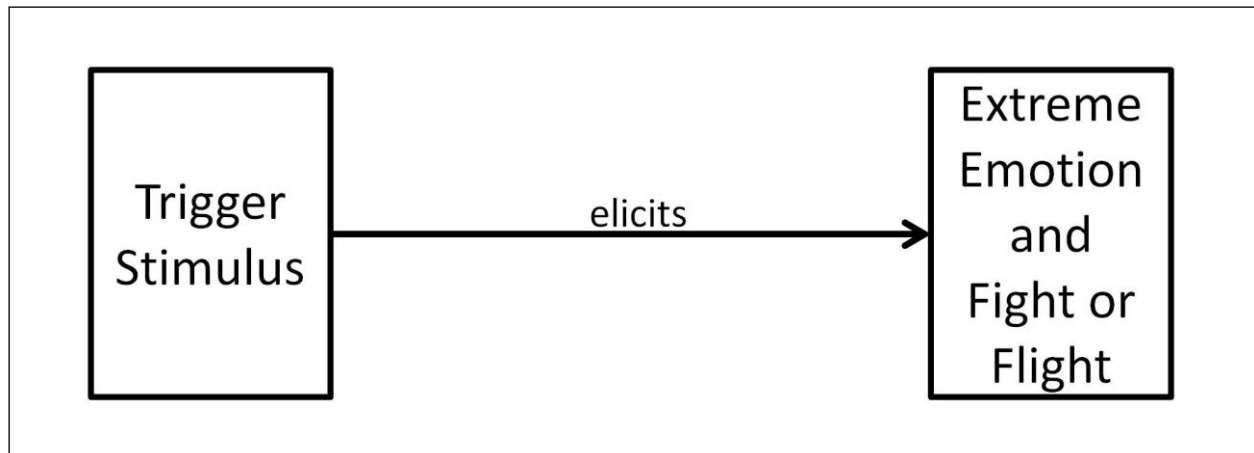


Figure 1. Misophonia as a one-step process. The trigger stimulus directly elicits extreme emotions and fight-or-flight responses.

2. Hypotheses

We propose the following hypotheses regarding the etiology, composition, development, and maintenance of misophonia. We believe that these hypotheses are true for the majority of individuals with misophonia, but we acknowledge that, in some cases, there may be other pathways and neurological processes that contribute to misophonia.

1. Misophonia develops as a Pavlovian-conditioned physical and/or emotional reflex.
2. Misophonia is often comprised of a physical muscle reflex that is elicited by the trigger stimulus and an emotional response that is elicited by the sensation of the physical muscle movement (see Figure 2). This proposes that misophonia is a two-step process. The person first perceives the trigger stimulus (i.e., hears the trigger sound) which elicits a physical muscle response – a purely physical

conditioned reflex. The physical sensation subsequently elicits the extreme emotional response.

3. The misophonic response to trigger stimuli is maintained (or strengthened) due to the physical reflex (e.g., muscle contraction) increasing in intensity during the conditioning episodes associated with each stimulus presentation. This increase in muscle contraction is associated with the emotional response elicited by the physical reflex (i.e., the emotional response of misophonia).
4. Additional trigger stimuli can develop by being paired with previously conditioned trigger stimuli and the reflex response they produce.

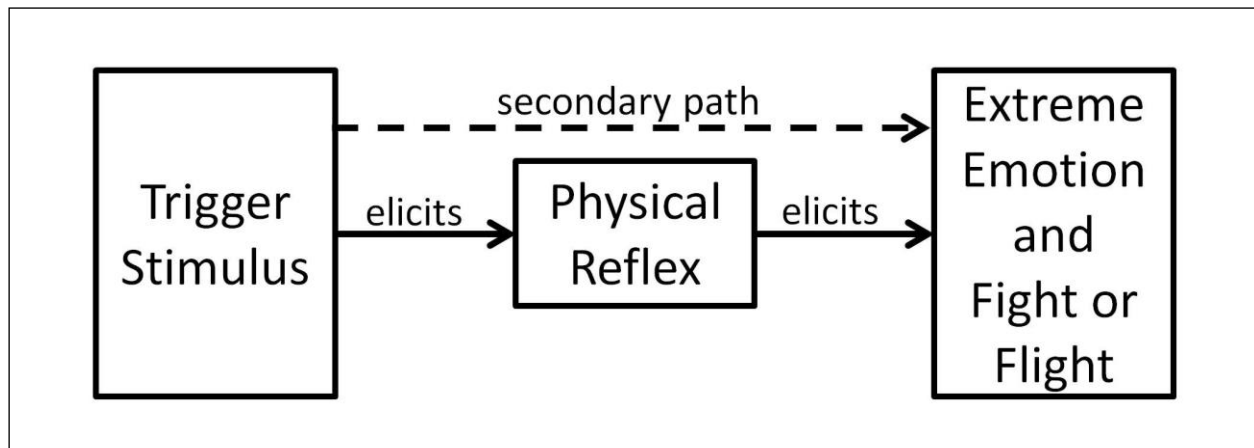


Figure 2. Misophonia as a two-step process. The trigger stimulus elicits a physical muscle reflex. Sensation from the physical muscle movement elicits the emotional response and fight-or-flight responses. In some individuals, a secondary path develops so that the trigger stimulus directly elicits emotions.

3. Discussion

3.1 Misophonia and Conditioned Reflexes

Jastreboff and Jastreboff (2014) stated that the misophonic trigger stimulus is part of a complex conditioned stimulus that includes the context and other factors. Although this was based on professional judgment, rather than a finding of an experimental study, they have treated individuals with misophonia for over a decade. They reported that of the 184 patients treated for misophonia, 83% showed significant improvement in their condition. Their treatment used several protocols designed to produce active extinction of the misophonic reflex. The success of their treatment method provides support for the hypothesis that misophonia is a Pavlovian-conditioned reflex.

The basic theory of Pavlovian conditioning is that pairing a neutral stimulus (NS) with an unconditioned stimulus (US) establishes an association that results in the NS producing a conditioned response similar to the unconditioned response (UR) elicited by the US. With misophonia, it is often difficult to identify an unconditioned (or conditioned) stimulus that was initially paired with the misophonic trigger stimulus (NS).

Research by Donahoe and Vegas (2004) indicated that the model of conditioning based on the idea that the pairing of the US and CS serves as the mechanism of the conditioning process is incorrect. Using an unconditioned reflex in pigeons with a half second stimulus-to-response delay allowed them to test the relationship between the NS and the US compared to the NS and the UR. Their research showed that it was the pairing of the NS and the UR that created conditioning, rather than the pairing of the NS and the US. All previous research on conditioning used a stimulus-response with no delay, making it impossible to determine if conditioning was due to the temporal relationship between the NS and the US or the NS and the UR.

The Donahoe and Vegas finding has been supported by research with neonates that develop a conditioned response in which the NS/CS was vanilla odor and the

UR/CR was a low state of arousal (Rattaz, Goubet, & Bullinger, 2005; Goubet, Strasbaugh, & Chesney, 2007). In these studies, there was no discrete US; the neonates were simply exposed to the vanilla for 11 or 16 hours, during which time they were calm or sleeping. After conditioning, the neonates maintained a lower state of arousal (i.e., did not cry) during a heel-stick when smelling the vanilla, and they cried less when the CS was presented after a heel-stick. We present the following cases to support classical conditioning as the etiology of misophonia.

Case 1: John develops his first trigger.

John, now a middle-aged adult with misophonia, recalled developing his first trigger. He shared a bedroom with his brother. John suffered from anxiety as a child. One night, he was unable to sleep. His brother had allergies and his breathing produced an audible nasal sound. After hours of hearing his brother breathe, John went to the couch and slept. From that night on, he was triggered whenever he heard his brother breathe. This provides support for the hypothesis that misophonia develops as a Pavlovian conditioned reflex. It is proposed that the nasal breathing sound became associated with the physiological response from the distress he experienced (i.e., specific contracted muscles) and/or the emotional distress experienced from anxiety, inability to sleep, and annoyance aroused by hearing the breathing sound. When the sound was heard later, it elicited the conditioned physical and/or emotional response.

Case 2: Carla develops misophonia.

Carla, age 10, presented at the clinic with a primary misophonia trigger of her brother chewing. She said that when she heard the trigger, she felt immediate rage but no physical response. Carla often had conflict with her brother at the dinner table. Her mother reported that when arguing, she would stand, extend both arms and demand

that her brother stop staring at her. This clearly was an operant behavior that included tight arm and leg muscles. In this setting, she also heard the sound of her brother's open mouth chewing. At the clinic, a low strength recorded trigger stimulus elicited a visible jerk in Carla's arms and shoulders. When asked what she felt, she reported feeling the contraction of muscles in her arms and legs, but no anger, rage, disgust, or weaker precursors of these emotions. It seems that the trigger stimulus elicited the contraction of the same muscles that were contracted when she was arguing with her brother, which supports the hypotheses that misophonia develops as a Pavlovian-conditioned reflex and that the initial reflex response to a trigger stimulus is a physical reflex.

Case 3: Connor develops misophonia.

Connor, age 24, presented for treatment of misophonia with severe auditory triggers of chewing, sneezing, mouth breathing, and smacking lips, and a visual trigger of someone touching their glasses. He developed misophonia while serving in the Marines in Afghanistan 2 years earlier. He reported that he also had a current diagnosis of PTSD. In Afghanistan, it was common to go on patrol as a squad, and upon returning to base, be in close quarters for eating. When tested for his initial physical misophonic reflex, his head visibly turned to the right, and he reported that he felt contraction of the muscles in his right arm including making a fist. The response was the same whether the trigger sound originated from on his left or right side. This response seems similar to orienting to a sound of danger on his right side. The misophonic triggers did not elicit PTSD responses.

Case 4: Bill develops misophonia.

Bill was in good health, in his early 30s, with no history of mental health problems. He presented with misophonia trigger stimuli of mockingbird chirps and lesser triggers to some other birds. One year earlier, mockingbirds had built their nest near Bill's bedroom window. Mockingbirds have a unique characteristic of singing 24 hours a day. The singing prevented Bill from sleeping and, over time, he developed a misophonic response to each of the five distinct calls of the mockingbird. Since then, he experienced a generalization of trigger stimuli to other (but not all) birds, though the elicited response to other bird chirps was less severe. Bill's physical reflex was a "chill" on his upper arm and a sensation on the sides of his head. This supports both the first and second hypotheses.

Case 5: Paul develops an aversive reflex response.

Consider the case of Paul, a middle-aged professional in good mental and physical health. He accepted a position in which he often received phone calls about problems he needed to handle. Paul developed a chest muscle contraction reflex to the default ringtone of the phone. It may be presumed that the chest muscle contraction was a physical response that accompanied the emotional reaction associated with the stress of the phone calls. He changed the ringtone to one which did not elicit the reflex; however, in time, the chest muscle contraction reflex developed to the new ringtone. He changed the ringtone several times, with the same result each time. Finally, he set his phone to vibrate only, and the reflex developed to the vibration ring of the phone. He also triggered to the ring of a phone on television, so it was clear that the sound elicited the reflex, independent of the caller or purpose of the call. In Paul's own words, "I hear the ring and my chest muscles jump, and I don't like it!" Paul's presenting problem was

limited to his irritation with the physical reflex. He did not experience any emotion similar to those accompanying the stressful phone calls. This reflex did not restrict or impair his activity in any way, but was still an aversive reflex to a typically occurring sound. We propose that any aversive muscle contraction reflex to sound or other stimuli could be termed a misophonic reflex.

These cases support the assertion that misophonia is an aversive conditioned reflex that develops when a person is in a state of distress and hears a repeating sound. In most of these cases, the sound could be a source or contributing factor for distress. In the case of Carla, it is not clear that the sound contributed to her distress, but the sound was being made by the person who was the source of her distress.

3.2. Misophonia as an aversive physical reflex

The defining characteristic of misophonia is the extreme emotions elicited by a commonly occurring soft sound (Schröder et al., 2013). Outwardly, misophonia appears to be a monolithic disorder that is virtually identical in each person. There is a common set of sounds (i.e., eating and breathing sounds) that are triggers and a common emotional response. The extreme emotions are described as involuntary and unavoidable, so it appears that the trigger directly elicits an emotional response. Each individual with whom we have worked reported that they tried to remain calm when exposed to triggers, but that it was impossible to do so. The physiological response of increased skin conductance identified by Edelstein et al. (2013) and increased general muscle tightness after repeated triggers validate the verbal report of increased arousal after misophonic triggers. From work with misophonic clients, we find that misophonia is not a monolithic disorder since there are unique manifestations in each individual. Although there is a common set of trigger sounds, the first trigger stimulus a person

develops is unique to that individual and is generally the sound made by a specific individual close to the misophonic person. We have also found that most individuals have a unique physical reflex to the trigger stimulus, although many are unaware of that physical reflex.

The extreme emotional response of misophonia seems to prevent recognition of the physical reflex. Indeed, in the case of Carla, while she was only aware of her rage response to the trigger stimulus, she had a visible physical reflex that was elicited by a reduced intensity recorded trigger stimulus. Some individuals confuse an operant response, such as covering their ears, with their reflex. Others have identified an accompanying physical response to anger, such as an accelerated heart rate or general muscle tension. One woman reported that her physical reflexes were tight arms, shoulders, neck and increased heart rate. When tested, they were only shoulder muscles and a “bump” of the heart. Since there was no increase in heart rate, it is likely she felt an elicited muscle contraction and not a change in autonomic arousal.

To allow the client to identify their physical misophonic reflex, the trigger stimulus is presented at a low volume and short duration in a controlled setting. This allows the individual to be aware of the bodily sensation and minimize the emotional response. This process is facilitated by recording the trigger stimulus so that volume and duration can be controlled. Identification of the physical reflex can be accomplished using a smartphone app (i.e., Misophonia Reflex Finder) designed for this purpose. We have found that over 95% of the individuals can identify an immediate physical reflex to the trigger stimulus, with only a few individuals reporting that regardless of how weak the stimulus or how calm they are, they only experience irritation. The reflexes are diverse. Patients have reported muscle contractions of shoulders, neck, whole arm, upper arm,

only the left upper arm, legs (in many variations), toes, abdomen, chest, jaw, hands open, hands making a fist, face, squinting, gasping and more. Other patients reported internal reflexes including stomach constriction, nausea, intestine constriction, esophagus constriction, sexual arousal, urge to urinate and unidentified movement sensations in the chest cavity.

3.3. Misophonia as a two-step reflex process

Treatment for misophonia has included a counterconditioning treatment known as the Neural Repatterning Technique (NRT) in which a low intensity, short duration trigger stimulus was intermittently provided during a continuous positive stimulus such as upbeat or relaxing music (Dozier, 2015). In such a situation, it was common (but not universal) for individuals to report that they felt a very mild physical reflex, but no accompanying emotional response, thus demonstrating that the physical reflex could occur independently of the emotional response. As shown in Figure 2, the initial response to a trigger is a physical reflex.

As noted above, misophonic individuals all reported that it had been impossible for them to not become upset when exposed to trigger stimuli, although they had repeatedly tried to remain calm. It is commonly said that “anger is a choice,” or that “another person cannot make you mad.” This may be true for verbal behavior. The meaning of the words is determined by a person’s learning experience with language and other social factors. A person’s response to a statement such as “I hate you” is affected by their evaluation of the context and social dynamics at that moment. With misophonia, however, anger is not a choice. There are two plausible constructs for emotional response being elicited by the initial physical reflex. The first construct is as follows. The sound elicits an intrusive, uncomfortable reflex response. We posit that this

physical reflex response is a form of physical assault on the person, although the actual physical assault is performed by their autonomic nervous system. The response to the repetitive physical reflex is extreme emotions (see Figure 2).

Aversive stimuli evoke fight or flight emotions in humans (Berkowitz et al., 1981; Berkowitz, 1983). The strength of these emotions is affected by a number of factors, and the instigation to aggression may not evidence itself in overt behavior. This is consistent with the emotions for misophonics, and the reported effort of misophonics to resist aggressive impulses. Furthermore, activity in the limbic system of humans in response to aversive odorants (Zald & Pardo, 1997) and to aversive gustatory stimuli (Zald, Lee, Fluegel, & Pardo, 1998) has been demonstrated. We argue that the aversive physical misophonic reflex evokes the commonly reported emotions of hate, anger, rage and disgust.

The second construct is that the physical misophonic reflex elicits a conditioned emotional response. The physical reflex is intrusive and difficult to not perceive; even when a person tries to use a technique to avoid attending to the auditory stimulus, the physical reflex is perceived and elicits the emotional response. This seems to be supported by the case of Martha, who had a very weak physical misophonic reflex.

Case 7: Martha eliminates her misophonia

Martha was a professional in her mid-forties with a lifelong history of misophonia ranging from mild to extremely debilitating. Efforts to decrease symptoms included extensive work to reduce autonomic reactivity, which included breath work, relaxation techniques, noise reduction headsets, and musician plugs. She had reduced her misophonia to the point that she rarely experienced extreme emotions, but she was still occasionally agitated by one trigger. After listening to a recording of the trigger in

preparation for the NRT treatment, she reported that she became aware of the muscles behind the ear contracting when she heard the sound. She used the NRT treatment (see section 3.6) for the trigger stimulus, and the reflex extinguished. Once the reflex extinguished, the real-life trigger stimulus no longer elicited negative emotions.

3.4. Respondent extinction and misophonia

If misophonia is a conditioned reflex, why does the elicited response to the trigger stimulus not die out – why does respondent extinction not occur? Traditional Pavlovian respondent (or active) extinction occurs when you remove the US and thereby also remove the UR. The CR is weaker than the UR, so the CS is paired with a progressively weaker and weaker CR, until the CR no longer occurs. The optimum time delay of the NS to the UR for conditioning of a skeletal muscle response is about half a second (Pierce & Cheney, 2013, p. 70). One study showed that for a conditioned eye-blink response, the optimum delay was 0.4 seconds for young adults, but 1.0 seconds for older adults (Solomon, Blanchard, Levine, Velazquez, & Groccia-Ellison, 1991). For an autonomic response, such as sweat secretion, a delay of 5 to 30 seconds is most effective (Pierce & Cheney, 2013, p. 70). Research on conditioning often uses very short CRs such as a dog's leg extension or an eye-blink so the muscle is at rest for most of the conditioning time, which is approximately 0-5 seconds after the stimulus (Catania, 2013, p. 246). With misophonia, individuals report that the physical reflex, such as shoulder or stomach constriction, is held for much longer than 5 seconds.

We propose that the immediate negative emotions following trigger stimuli and multiple presentations of the trigger stimuli cause an increased contraction in the muscles that are constricted by the initial physical misophonic reflex. Anger is typically accompanied by increased tension in skeletal muscles, and multiple presentations of the

stimuli could cause progressively elevated contraction of the muscle. This would pair the trigger stimulus with a stronger physical action and could serve to strengthen the conditioned physical reflex response. Thus, a self-strengthening situation is created, wherein exposure to triggers increases or maintains the severity of the physical misophonic reflex; hence, active extinction (respondent extinction) cannot occur. In fact, consistent with the progressive worsening of misophonia, it is anecdotally reported that exposure therapy (i.e., listening to the therapist make the trigger sound) and attempting to tolerate misophonic triggers in real life generally make misophonia worse.

3.5. Acquiring new misophonia trigger stimuli

We propose that additional trigger stimuli can develop by being paired with previously conditioned trigger stimuli and the reflex response they produce. Simultaneous conditioning required a temporal relationship between the NS and the US/UR whereby the NS precedes the US/UR by 0-5 seconds (Catania, 2013, p. 246). This relationship occurs when a visual stimulus (e.g., chip into mouth) precedes an established trigger stimulus (e.g., crunch). As expected, this visual stimulus often becomes a trigger stimulus. Because triggers are strengthened through repeated exposure, the visual of “chip into mouth” can remain a trigger even if it is no longer paired with the auditory (crunch) trigger.

3.5.1 Case 6: Conditioning a new trigger stimulus

Brent, a middle-aged man, had several visual triggers. He reported that his physical response to triggers was a constriction of his intestines. We attempted treatment of a visual trigger with the Visual Trigger Tamer app. Because he was using music for the positive stimulus, a chime was included prior to the trigger so he would know when to view the trigger video. He was cautioned to keep the trigger stimulus

short so his misophonic response would be weak and brief. Obtaining a brief response was particularly difficult because the intestine constriction would persist if the trigger was too strong. We also hoped this would minimize the risk of the chime becoming a trigger. Brent reported that in an effort to speed the treatment effect, he increased the trigger strength. When he did this, the chime began to elicit intestine constriction and so the treatment was halted. This demonstrated developing a new trigger stimulus through Pavlovian conditioning.

3.5.2 Other ways to develop triggers

Since the physical misophonic response is sustained for many seconds or minutes, any other repetitive stimulus, including in other sensory modalities, occurring simultaneously with the misophonic response may become a CS for that response due to its pairing. These could include the jaw movement of chewing, variations of the trigger stimulus such as the eating sounds of a non-trigger person or the click of a fork on a plate.

A third way that a new trigger can be acquired is to develop it in the same manner as the original trigger, by being in a state of physiological distress (i.e., tense muscles) and hearing a repeating sound. The case of Aubrey (age 37) fits this description. When she heard whistling or a pen click, her neck and shoulders flinched, but when she heard coughing or snorting, her esophagus constricted. Since the physical misophonic reflexes were different, it appears that the misophonic triggers-reflexes were developed independently.

3.6. A treatment for misophonia as a conditioned reflex

If a treatment based on the proposed model is effective, it would lend support to the hypothesized model. As it is posited that the emotional response following trigger

stimuli causes the reflex to strengthen, a treatment that exposes a person to stimuli under conditions in which the emotional response is not elicited should weaken the reflex. A counterconditioning treatment called the Neural Repatterning Technique (NRT) was developed to provide exposure to the trigger stimulus while not eliciting the emotional response (Dozier, 2014; Dozier, 2015). A recorded trigger stimulus was presented with reduced volume, duration, and frequency of the trigger. The counterconditioning stimulus was selected by the subject to elicit a pleasant experience. Based on immediate feedback from the patient, the trigger stimulus was adjusted so that the patient's physical response was brief and weak. Under these conditions, the misophonic physical response would be expected to demonstrate respondent extinction or counterconditioning.

Case 8: Karen and the neural repatterning technique treatment

Karen was a 48-year-old woman in good physical and mental health. Her most distressing misophonic triggers were generated by her husband. She described her experience as feeling as though "someone is running a shovel through the center of my chest and out my back." However, she noted that this was a somewhat metaphorical description of a combined emotional and physical response. Karen reported that she had chronic asthma during the time she developed her first misophonia trigger, and her physical misophonic reflex was similar to the way she would strain to breathe. This consisted of pulling her shoulders forward and upward, contracting her face muscles, and straining to breathe in. It is plausible that Karen's misophonia developed as a conditioned physical reflex to the trigger stimulus.

Treatment was provided via internet telecommunications, with audio files provided for homework. The positive stimuli in live sessions were conversations about positive

career events and other happy topics. The positive stimulus for homework to be completed independently was upbeat music or listening to the discussion recorded in a live session. Karen averaged 4 homework sessions per week. She responded well to the counterconditioning treatment and the physical reflex to the trigger stimulus extinguished in the treatment setting. When she heard the live trigger stimulus, she had a small physical reflex response, which she could often ignore, while the emotional response typically did not occur. In her case, greatly reducing the physical reflex through counterconditioning also eliminated or greatly reduced the emotional response to the trigger stimulus. This provides further support for the hypothesis that the emotion of misophonia is elicited by the aversive physical reflex because the counterconditioning treatment was aimed at the physical misophonic reflex. The treatment took 2 weeks each for three auditory triggers and 9 weeks for a single visual trigger. At the 4- and 10-month posttreatment follow-up assessments, Karen reported that her responses to three of her triggers were the same as at the end of treatment, and one had strengthened slightly. In addition, she stated that her overall concern about her misophonia was minimal and declining, based on her misophonia severity assessment scores (Dozier, 2015).

Subsequent experience suggests that the substantial improvement Karen enjoyed is likely to occur only in cases wherein triggers are restricted to the sound of a single individual or a single setting. When the treatment has been used on a person with a generalized trigger (e.g., a sniff of anyone, anywhere), the reflex extinguished in the treatment setting with a recorded trigger, but effects did not generalize to real-world trigger exposures. Treatment benefits seem to be limited to individuals who have

specific trigger sounds from specific individuals. Misophonia often involves more limited triggers that are more responsive to this treatment approach when first developing.

Case 9: Mary and NRT treatment

Mary, a university student home for the summer, presented with two troubling triggers produced by her mother. She reported having severe misophonia at university, but those triggers did not occur at home. Her misophonia severity score was 49. When tested, she identified her physical reflex as “making a fist.” She practiced Progressive Muscle Relaxation twice a day for a week before starting the NRT treatment, and then once a day during treatment. Her positive stimulus for the NRT treatment was upbeat music. Her treatment sessions were conducted entirely as independent homework using the Misophonia Trigger Tamer app.

The first trigger sound was a spoon scraping on a bowl. She completed 5 treatment sessions over a 2-week period, with a total of 368 exposures to this trigger. Her second trigger was the sound of her mother chewing. She completed 4 treatment sessions over a 1-week period, with a total of 178 exposures to this trigger. During the treatment sessions, she tried to keep her hands relaxed at all times. She reported that, during treatment, she felt the contractions of her muscles when exposed to a trigger, but had no emotional response. She also worked on relaxing her misophonia reflex by wiggling her fingers whenever she was exposed to a live trigger. After 3 weeks of treatment, she reported that she no longer had any misophonic response to the spoon on bowl sound, and she had a weak response to live chewing sounds but none to the recorded sounds.

She completed the misophonia severity questionnaire 21 days after her first treatment and her score was 13, indicating that her misophonia rating had decreased from severe to mild. Perhaps her quick response to this treatment was aided by relaxing

her fist muscles, which could have accelerated the respondent extinction process. At a 6-month follow-up, her misophonia severity score was 7.

Case 10: Vira and emotions with NRT treatment

The presentation of misophonia is variable and some individuals report only a physical response to weak triggers during treatment while others also describe the emotional response. Vira, age 74, developed misophonia as a young child. She had a strong physical reflex and extreme emotions to live triggers produced by family members. She did not have any triggers produced by strangers. During the NRT treatment sessions, she adjusted the trigger settings to keep the emotional response at a low level, but she also felt the physical reflex, which seemed less severe than her emotional response.

Vira responded positively to the NRT treatment, which was performed with the Misophonia Trigger Tamer and Visual Trigger Tamer apps over an 8-month timespan. The Misophonia Trigger Tamer was used to treat auditory triggers and the Visual Trigger Tamer was used to treat visual or visual/auditory triggers of gum chewing. The treatment app provides precise control of the volume and duration of the triggers. These parameters are adjusted by the patient to maintain a brief, weak response to the trigger. Slight increases in these parameters to maintain the reflex response indicate a reduction in the misophonic response to the trigger. Throughout most of the treatment, Vira experienced both an emotional and a physical response, and the misophonic response steadily declined to each trigger stimulus. For the last month of treatment, Vira experienced only an emotional response to the trigger. After a month of no progress with this trigger, treatment was terminated. Her misophonia severity score dropped from 46 at the start of treatment to 4 at the end of treatment. The treatment eliminated her

real-world physical and emotional response to all trigger stimuli that she routinely experienced. In Vira's case, it seems that that a misophonic stimulus directly triggered a conditioned emotional response, and the conditioned emotional response did not extinguish with the counterconditioning treatment.

Figure 2 shows the development of a direct emotional response to the trigger stimulus as a secondary process. We view the directly elicited emotional response as a secondary process because of the many cases in which there is a clear physical reflex which occurs independently of the emotional response during the NRT treatment. Furthermore, in Vera's case, as long as the NRT sessions produced a physical reflex, the response in treatment extinguished and her real-world physical and emotional responses abated.

Another individual, who did not report an immediate physical reflex to the trigger stimulus, attempted the NRT treatment. Even with multiple experiences of the trigger at a low level, he only felt an emotional response to the trigger stimulus, and the treatment had no effect on his misophonia.

3.7. Counterconditioning by blocking the reflex

Some misophonia physical reflexes can be reduced, halted, or blocked by a willful action or a competing response. For example, if the reflex is a gasp, then slowly breathing in and out may reduce the strength of the reflex response. Mary (see case 9) reported in the follow-up discussion that she relaxed and sometimes shook her hands to relax them when repeatedly triggered. She reported that this has greatly reduced her emotional response to real-life triggers that were not treated with the NRT treatment. John (see case 1) had misophonia for over 30 years and was very skilled at muscle relaxation. He reported that, for many years, he controlled his anger response after a

trigger by relaxing muscles throughout his body. He also reported that he eliminated all of his triggers over several months, one by one, by relaxing his muscles before a trigger, such as when he was in an environment with a repeating trigger. His initial physical misophonic reflex was pulling his shoulders toward his ears. In his case, eliminating his physical misophonic reflex eliminated his misophonia. This also provides support for misophonia as a two-step process because John only worked on his physical reflex response to the triggers, and reported that this completely eliminated his emotional misophonic response.

3.7.1 Counterconditioning a sexual arousal reflex

Two individuals received a single session counterconditioning treatment that used a tickle reflex to block a sexual arousal reflex. In the first case, a trigger produced by the therapist was used and, in the second case, a recorded trigger and the Trigger Tamer app were utilized. In both cases, a cohort tickled the patient as soon as the cohort heard the stimulus. This completely blocked the arousal reflex. The trigger stimulus was emitted every 30 seconds for 20 minutes. Tickling was omitted for the first and the last triggers to observe the strength of the misophonic reflex. In both cases, sexual arousal occurred for the first trigger stimulus, but there was no sexual arousal for the final trigger stimulus.

It is acknowledged that these are based on verbal reports, so there is some concern about reliability. For the first case, the therapist conducted the session and solicited immediate verbal reports from the patient. For the second case, the patient and cohort conducted the session and reported to the therapist by email. Typically, hearing a trigger every 30 seconds for 20 minutes would lead to increased distress with no reduction in the triggered reflex response. In these cases, the trigger did not elicit the

reflex. It is plausible that the reflex did not occur because of counterconditioning or extinction of the conditioned reflex. These cases support that misophonia includes a conditioned physical response to a stimulus as well as the two-step view of misophonia shown in Figure 2.

3.8. Spontaneous Recovery of Triggers During Treatment

Spontaneous recovery is a well-documented phenomenon that is observed when a conditioned reflex extinguishes after multiple presentations of a CS without the US, and then the conditioned reflex recovers slightly with the passage of time (Catania, 2013). Consistent with spontaneous recovery, patients have reported a strengthening of the misophonic response to a trigger between NRT treatment sessions. The Trigger Tamer app allows the volume and duration of the trigger to be set precisely. Vera (case 10) reported that she needed to start a treatment session with the trigger volume slightly below the final volume of the previous treatment, to achieve the desired low level misophonic response. Although there are other plausible explanations, the effect Vera observed could have been spontaneous recovery of a conditioned reflex.

4. Conclusion

4.2. Hypotheses support

Each of the first five cases discussed above detail the pairing of a repeating auditory stimulus with a physiological and emotional response during the development of misophonia. Although these are anecdotal reports, each describe a pairing process that could lead to learning a conditioned response to an auditory stimulus. Case 1 is unusual because John remembered the first day he had a misophonic trigger response,

and the night before when he heard the trigger stimulus while in a state of distress.

Case 2 provides support for the etiology of misophonia as a conditioned physical reflex response (i.e., arm, shoulder and leg muscle contraction) to a specific trigger stimulus (i.e., brother's chewing sounds). It is noteworthy that this physical response was part of an operant behavior (i.e., verbal behavior), and not a response elicited by a US or CS.

Case 3 provides an example of a physical reflex response that could have been commonly elicited by the danger that Connor regularly experienced. Case 4 and 5 are particularly supportive because they illustrate the recent development of a physical response to a trigger stimulus in healthy adults. Case 8 also supports the etiology of misophonia as a conditioned physical reflex, wherein the individual simultaneously experienced the muscle strain of trying to breathe and the auditory stimulus that became the first trigger.

These six cases provide support for hypothesis 1, which states that misophonia develops as a physical and/or emotional reflex through classical conditioning. The simplest explanation of the mechanism for the development of a physical reflex to a specific stimulus is Pavlovian or classical conditioning. The possibility that the reflex develops due to maturation would be contraindicated by the wide variation in the ages at which misophonia is first observed (Edelstein et al., 2013). The occurrence of a change in the strength of the misophonic response in treatment that is consistent with spontaneous recovery further supports this hypothesis. Additionally, the distinct physical reflex to the stimulus in cases 2, 3, 4, and 5 also provides support for the idea that misophonia can be a conditioned physical reflex.

Hypothesis 2, that misophonia consists of an initial physical muscle reflex elicited by the trigger stimulus and an emotional response elicited by the sensation of the

physical muscle movement, is clearly supported by the observation that a large majority of the individuals with whom we have worked clearly identified a physical reflex response to the trigger stimulus that was independent of the negative emotional response. The independent nature of the physical reflex is shown by the patients' reports of experiencing only the physical reflex response to the trigger during the NRT treatment. If misophonia was primarily an emotional response to the trigger stimulus, then this emotional response would occur in response to a low intensity trigger stimulus, or a combination of emotional and physical reflex responses simultaneously elicited by a low-intensity trigger stimulus. Furthermore, when the physical muscle response to the trigger is eliminated during the NRT treatment, the response to a real-world trigger has a greatly reduced physical muscle and emotional response, indicating that the emotional response is in proportion to the physical response. Finally, when John (case 1) relaxed his muscles during trigger exposure, it eliminated his misophonic response to that trigger, which indicates that his misophonia was primarily a physical reflex which elicited the emotional response. We believe that the wide variety of physical misophonic reflexes reported compared to the limited number of misophonic emotional responses also supports hypothesis 2.

Hypothesis 3, that the misophonic reflex is maintained or strengthened by the emotional response to the trigger stimulus, is supported by the decline of the misophonic response with the NRT treatment. In real-life settings, the misophonic response generally maintains or strengthens with time, and individuals with misophonia regularly experience extreme emotions in response to trigger stimuli. With the NRT treatment, the patients have minimal or no emotional response to trigger stimuli and the physical reflex decays. Sometimes it decays very quickly, and sometimes it decays very slowly.

Case 6 provides specific support for the hypothesis that new trigger stimuli can develop by being paired with previously conditioned trigger stimuli. Developing conditioned stimuli by pairing an NS with a CS is a basic principle of classical conditioning. If misophonia is a conditioned reflex, then when an NS paired with a trigger stimulus (the CS), the NS can become a new trigger stimulus, as observed in case 6.

4.3. Neurological functioning

Some consider misophonia to be caused by malfunctioning or defective neurology in higher brain structures (Møller, 2011). We view misophonia as a conditioned response that develops because the neurology is functioning normally. Once the conditioned response develops, it strengthens with repeated exposure and generalizes to other sounds and sights in the environment. Misophonia may be one of the simplest of all psychological conditions, i.e., an aversive physical reflex which is maintained or strengthened by a physical response to the negative emotions it creates.

4.4. Impact of misophonia

The impact of misophonia on the life of the individual varies greatly, ranging from debilitating to having almost no impact on their quality of life. One factor that affects the severity of misophonia is the reactivity of the individual to aversive stimuli. Cases using CBT therapy have successfully remediated misophonia by reducing reactivity (Bernstein et al., 2013; McGuire, Wu, & Storch, in press).

The impact of misophonia, in large part, seems to depend on the quantity, pervasiveness, and ability to avoid trigger stimuli. If a person is rarely exposed to trigger stimuli, misophonia has virtually no impact on their life. But if triggers develop to stimuli

produced by close family members (e.g., son or daughter), the trigger may be inescapable, cause great distress to the individual, and impact their quality of life. One individual with whom we worked only had one trigger, but it was commonly heard in university classrooms, and so created a very severe case of misophonia for this student. Indeed, the first trigger stimulus an individual develops may be the start of the chronic, debilitating condition of misophonia, or it may be a single reflex that is simply annoying.

In children, misophonia is often misdiagnosed by parents and professionals as noncompliance, hypersensitivity, or unreasonable emotional outbursts. Considering the finding that 20% of study participants had clinically significant misophonia (Wu et al., 2014), there is a great need for increased awareness of misophonia by parents, doctors, psychiatrists, therapists, and school counselors. Having a sound evoke an aversive physical reflex and extreme emotions is incomprehensible to most parents, spouses, and other close individuals, but when the trigger sounds generalize to work and school settings, normal functioning in those environments becomes impaired and, in some cases, impossible. The extreme, emotional response to the trigger sounds often results in self-isolation to avoid the trigger stimuli; thus, the misophonic emotions and overt responses to triggers can have a significantly negative impact on interpersonal relationships.

4.4. Misophonia or Conditioned Aversive Reflex Disorder

The term “misophonia” has several shortcomings in terms of describing this condition. Firstly, it only refers to the emotional response to auditory triggers; many individuals have visual triggers and there are anecdotal reports of olfactory and tactile triggers. This being the case, Schröder (2013) proposed that an extreme emotional response to visual triggers be named misokinesia, which would be a second name for

the same phenomenon. Secondly, the term misophonia puts the focus on the individual's experience of a strong dislike of sound. Liking or hating something is generally an evaluative process that can be altered by thoughtful consideration, but this is generally not the case with this condition. Finally, a disorder that is fundamentally an emotional response places the neurological emphasis on the limbic system while a disorder that is fundamentally a conditioned reflex response places the neurological emphasis on the autonomic nervous system. This distinction has important implications for both research and the development of treatments. Therefore, the term "misophonia" does not clearly indicate the reflexive nature of the condition, which can also be a source of misunderstanding when communicating to family members, teachers, and employers.

We propose that Conditioned Aversive Reflex Disorder (CARD) is a more descriptive and appropriate name for this condition. It puts the focus on the reflex nature of the disorder and on the etiology of reflex, which is classical conditioning. CARD easily incorporates all modalities of trigger stimuli. Specifying it as a disorder requires a diagnostic criteria to determine a clinical level vs. nonclinical level of such reflexes. This disorder presents with great variety of aversive reflexes, where one person may have a single aversive reflex, as illustrated above in the case of Paul, while another individual may suffer a debilitating condition that causes them to be unable to tolerate a typical work environment.

4.5 Final remarks

Currently, no treatments have been evaluated in a controlled study and shown to reliably relieve misophonic symptoms. Controlled studies of available treatments are needed, as well as research of all aspects of misophonia (or CARD). The proposed

view of misophonia as a conditioned physical reflex that elicits an emotional response needs to be independently evaluated. If true, this will have significant implications for misophonia research and treatments. Misophonia has historically been the domain of audiologists, and now seems to be garnering more attention as a neurological or psychological condition. Research needs to include all disciplines, especially those professionals who are expert in eliminating or altering conditioned reflexes.

References

- Berkowitz, L. (1983). Aversively stimulated aggression: Some parallels and differences in research with animals and humans. *American Psychologist*, 38(11), 1135-1144. doi:10.1037/0003-066X.38.11.1135
- Berkowitz, L., Cochran, S. T., & Embree, M. C. (1981). Physical pain and the goal of aversively stimulated aggression. *Journal of Personality and Social Psychology*, 40(4), 687 -700. doi:10.1037/0022-3514.40.4.687
- Bernstein, R. E., Angell, K. L., & Dehle, C. M. (2013). A brief course of cognitive behavioural therapy for the treatment of misophonia: A case example. *The Cognitive Behaviour Therapist*, 6(10), 1-13. doi:10.1017/S1754470X13000172
- Catania, A. C. (2013). *Learning* (5th ed.). Cornwall-on-Hudson, NY: Sloan Publishing.
- Donahoe, J. W., & Vegas, R. (2004). Pavlovian conditioning: The CS-UR relation. *Journal of Experimental Psychology: Animal Behavior Processes* 30(1), 17-33. doi:10.1037/0097-7403.30.1.17
- Dozier, T. H. (2014, February). *Misophonia: An aversive conditioned reflex to soft sounds*. Poster presented at the annual convention of the California Association for

Behavior Analysis, San Francisco, CA. Available from:

<http://misophoniatreatment.com/wp-content/uploads/2014/03/Misophonia-as-a-conditioned-reflex-2014-CalABA.pdf>

Dozier, T. H. (2015). Counterconditioning treatment for misophonia. *Clinical Case Studies*. Published online before print January 20, 2015. doi: 10.1177/1534650114566924

Edelstein, M., Brand, D., Rouw, R., & Ramachandran, V. S. (2013). Misophonia: Physiological investigations and case descriptions. *Frontiers in Human Neuroscience* 7. doi: 10.3389/fnhum.2013.00296

Goubet, N., Strasbaugh, K., & Chesney, J. (2007). Familiarity breeds content? Soothing effect of a familiar odor on full-term newborns. *Journal of Developmental & Behavioral Pediatrics* 28(3), 189-194. doi:10.1097/dbp.0b013e31802d0b8d

Jastreboff, M. M., & Jastreboff, P. J. (2002). Decreased sound tolerance and tinnitus retraining therapy (TRT). *Australian and New Zealand Journal of Audiology* 24(2), 74-84. doi:10.1375/audi.24.2.74.3110

Jastreboff, M. M., & Jastreboff, P. J. (2013). Using TRT to treat hyperacusis, misophonia and phonophobia. *ENT Audiology News* 21(6), 88-90.

Jastreboff, M.M., & Jastreboff, P.J. (2014). Treatments for decreased sound tolerance (hyperacusis and misophonia). *Seminars in Hearing* 35(2), 105-120. doi: 10.1055/s-0034-1372527

McGuire, J. F., Wu, M. S., & Storch, E. A. (in press). Cognitive behavioral therapy for two youth with misophonia. *Journal of Clinical Psychiatry*.

- Møller, A. R. (2011). Misophonia, phonophobia, and “exploding head” syndrome. In *Textbook of Tinnitus*, Møller, AR, Langguth, B, DeRidder, D, & Kleinjung, T (Eds.). NY: Springer Publishing.
- Pierce, W. D., & Cheney, C. D. (2013). *Behavior analysis and learning*. NY: Psychology Press.
- Rattaz, C., Goubet, N., & Bullinger, A. (2005). The calming effect of a familiar odor on full-term newborns. *Journal of Developmental & Behavioral Pediatrics* 26(2), 86-92. doi:10.1097/00004703-200504000-00003
- Schröder, A., Vulink, N., & Denys, S. (2013). Misophonia: Diagnostic criteria for a new psychiatric disorder. *PLoS ONE* 8(1), e54706.
- Solomon, P. R., Blanchard, S., Levine, E., Velazquez, E., & Groccia-Ellison, M. (1991). Attenuation of age-related conditioning deficits in humans by extension of the interstimulus interval. *Psychology and Aging* 6(1), 36-42. doi: 10.1037/0882-7974.6.1.36
- Zald, D. H., Lee, J. T., Fluegel, K. W., & Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain*, 121(6), 1143-1154.
- Zald, D. H., & Pardo, J. V. (1997). Emotion, olfaction, and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proceedings of the National Academy of Sciences*, 9(8), 4119-24.
- Wu, M. S., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2014). Misophonia: Incidence, phenomenology, and clinical correlates in an undergraduate student sample. *Journal of Clinical Psychology*, 70(10), 1-14. doi: 10.1002/jclp.22098