Title

An examination of contextual factors that influence auditory processing in misophonia and absolute pitch

Author
Edelstein, Miren Hope

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

An examination of contextual factors that influence auditory processing in misophonia and absolute pitch

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Experimental Psychology

by

Miren Hope Edelstein

Committee in Charge:

Professor Diana Deutsch, Co-chair
Professor V.S. Ramachandran, Co-chair
Professor Don Macleod
Professor Miller Puckette
Professor Steve Schick

2019

The Dissertation of Miren Hope Edelstein is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-chair

Co-chair

University of California San Diego

2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors, Dr. V.S. Ramachandran and Dr. Diana Deutsch for their invaluable mentorship throughout my graduate school career. I am extremely privileged to have worked with two individuals who are absolute giants in their respective fields and will carry with me what I have learned from both of them for the rest of my life.

I would also like to thank my cohort, most of whom have already graduated and moved on to successful careers in academia and industry, for their camaraderie throughout the long journey that was graduate school. From my cohort, I would like to specifically thank Dr. Mary Smith and Dr. Camille Toarmino, who became some of my closest friends and have always been supportive of me in all that I do. Additionally, I would like to thank my partner, Bradley Monk, who, for the past six and a half years, has helped guide me over many hurdles in my life (including research related ones), has kept me laughing and has lifted me up throughout this process, all while completing this program with me.

I am thankful for the support of my lab mates past and present (Dr. David Brang, Dr. Laura Case, Dr. Kevin Dooley, Zeve Marcus, Dr. Chaipat Chunharas, Nick Root and Liz Seckel) and for the support of my long-time collaborator from the University of Amsterdam, Dr. Romke Rouw, for her guidance and for believing in our work on misophonia from the very beginning. I would also like to recognize my former research assistants (Ronald Robertson, Diana Hereld, Jane Yang and Snehal Lochan) as well as all of the individuals who participated in these research studies. This body of work would not exist without their invaluable contributions in the lab. I would also like to acknowledge the psychology department’s graduate coordinators past and present (Rachael Lapidis, Carson Dance and Samantha Llanos), who tirelessly helped me navigate the program and always had the best interests of us graduate students in mind.

I would like to thank my parents for instilling in me a lifelong love of science and the arts, and for being supportive of my choice to pursue both. Graduate school has been a difficult journey and I am grateful that they never lost hope in my ability to finish. Lastly, I would like to recognize my close friends and bandmates (Vivian Feig, Vicky Le, Angeline Feng, Tina Yokoyama, Dr. Sherry Chak, Summer Swee-Singh, Dr. Joyce Cheng, Dr. Ashley Juavinett, Dr. Ethan McBride and Dr. Landon Klein). I am truly fortunate to have such remarkable people in my corner.

Chapter 1, in full, is a reprint of the material as it appears in *Frontiers in Human Neuroscience*. Edelstein, Miren; Brang, David; Rouw, Romke; Ramachandran, V.S. 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Edelstein, Miren; Monk, Bradley; Rouw, Romke; Ramachandran, V.S. The dissertation author was the primary investigator and author of this paper.

Chapter 3, is coauthored with Monk, Bradley; Henthorn, Trevor and Deutsch, Diana. The dissertation author was the primary investigator and author of this paper.

VITA

EDUCATION

Ph.D. in Experimental Psychology 2019
University of California, San Diego

M.A in Psychology 2012
University of California, San Diego

B.A. in Psychology (honors), minor in Music 2011
University of California, Berkeley

PUBLICATIONS

Brout, J.J., **Edelstein, M.**, Erfanian, M., Mannino, M., Miller, L., Rouw, R., Kumar, S., Rosenthal, M.Z. (2018). Investigating misophonia: a review of the empirical literature, clinical implications and a research agenda. *Frontiers in Neuroscience.* 12, 36

Edelstein, M., Brang, D., Rouw, R., Ramachandran, V.S. (2013). Misophonia: physiological investigations and case descriptions. *Frontiers in Human Neuroscience*. 7, 296

ABSTRACT OF THE DISSERTATION

An examination of contextual factors that influence auditory processing in misophonia and absolute pitch

by

Miren Hope Edelstein

Doctor of Philosophy in Experimental Psychology

University of California San Diego, 2019

Professor Diana Deutsch, Co-chair
Professor V.S. Ramachandran, Co-chair

This dissertation covers two unrelated topics related to human auditory processing: misophonia and absolute pitch (AP). Misophonia is a newly researched condition in which certain sounds evoke extreme distress, significantly impacting the quality of life in those who suffer from it. Absolute pitch, also known as "perfect pitch," is the rare ability to identify or produce musical pitches in isolation without the aid of a reference pitch. Absolute pitch is extremely rare, even among lifelong musicians. Although they are unrelated, these two groups do share a common thread: they both have highly specific associations with and responses towards particular sounds that are not seen in the general population. Chapter 1 of this dissertation provided the first empirical research study ever conducted on misophonia. This study characterized the symptoms of what was, at the time of publication, a largely unknown condition. Chapter 2 further details the misophonic condition, with a particular focus on the interplay between sound and contextual information. We systematically manipulated the information paired with certain sounds and discovered that the very same sound could be reported as significantly more or less aversive, by the same individual, within the same experimental session. Chapter 3 covers a study that examined how the performance of absolute pitch possessors on a pitch labeling task could be influenced by note timbre and instrument expertise. Findings revealed a congruency effect in which participants performed significantly better on the task when trial timbres matched their instrument of expertise and worse when trial timbres did not match their instrument of expertise, highlighting an interaction of factors that can produce variation in absolute pitch ability. Taken together, the studies in this dissertation further our understanding of how auditory stimuli are processed and linked with contextual information, and ultimately show how the information associated with certain sounds can affect how we respond to them.

GENERAL INTRODUCTION

This dissertation covers two interesting psychological topics related to human processing of auditory stimuli, and the studies herein examine two unrelated groups of individuals who have formed strong associations with specific sounds. The first two chapters are focused on a case population of individuals who suffer from misophonia, a condition in which certain sounds can evoke extreme distress, to the point where quality of life can become greatly impacted. The third chapter is a study involving musicians with a rare ability known as absolute pitch (AP), whereby these individuals are able to identify isolated pitches (e.g. musical notes) without requiring a prior tonal reference. Together these two topics dovetail as an exploration into how two distinct human populations process auditory stimuli. The results of these studies provide a unique perspective into the manner with which sound is coupled with contextual information, and the bearing of these associations on cognitive interpretations and physiological responses.

Misophonia is a relatively unexplored condition in which specific sounds cause an aversive response in individuals, characterized by negative emotions so intense that they are analogous to a fight-or-flight response (Jastreboff & Jastreboff, 2001; Edelstein et al., 2013; Schröder et al., 2013; Rouw & Erfanian, 2017; Potgieter et al., 2019). Common misophonic "trigger" sounds tend to be ordinary eating sounds, or repetitive sounds like pen clicking or keyboard typing (Edelstein et al., 2013; Schröder et al., 2013; Wu et al., 2014). Individuals with this condition report an acute sense of anxiety, panic, rage or even disgust when exposed to these trigger sounds (Edelstein et al., 2013; Wu et al., 2014; Kumar et al., 2017). Chapter 1 of this dissertation is a 2013 paper that is considered a seminal work on misophonia, as it was the first study to utilize scientifically rigorous methods to examine the condition and made early contributions to its scientific characterization. The first part of this study involved interviewing individuals with misophonia, through which we were able to gain a detailed overview of its symptoms, which aligned with proposed diagnostic criteria reported earlier that year by Schröder et al. (2013), and also shed light on several curious aspects of the condition (which later served as the inspiration for the study in Chapter 2). Additionally, the main experimental finding of this study showed that misophonic individuals not only rated auditory stimuli as more aversive than purely visual stimuli, but also exhibited a significantly heightened skin conductance response (SCR) for auditory as opposed to visual stimuli. As a result, this study was the first to experimentally show that individuals with misophonia experience heightened autonomic nervous system arousal to sounds, a response not seen to the nearly same extent in control participants.

As mentioned above, there were several curious characteristics of misophonia that warranted further exploration. One of these characteristics was the specificity of the misophonic response. Although there are always exceptions, many misophonic individuals indicate that they are particularly averse to trigger sounds produced by specific people and are often not triggered (or are triggered to a lesser extent) when these sounds are self-produced, produced by an animal/pet or produced by strangers (Edelstein et al., 2013). Frequently, the people whose sounds are the most triggering are friends and family members who are close with the individual with misophonia (Bernstein et al., 2013; Edelstein et al., 2013; McGuire et al., 2015). This finding suggests that a misophonic aversive stimulus often consists of more than just the low-level features of a sound and must involve some learned or conditioned contextual cues (Jastreboff & Jastreboff, 2001; Bruxner, 2016) which together with the sound generate a Gestalt, auditory-evoked trigger. As a follow up to the study in Chapter 1, the study in Chapter 2 explored this characteristic of misophonia in detail through experiments that utilized traditional misophonic triggers (e.g. human eating sounds) as well as similar sounding stimuli (e.g. animal eating sounds, non eating sounds such as snow crunching etc.); importantly the sounds were presented with varying contextual information. One major finding from this study was that an individual could find the same sound, when it was encountered again, to be significantly more or less aversive depending on the amount of and type of contextual information that was presented along with it.

Chapter 3 switches gears and examines musicians, specifically pianists and violinists, with AP. AP ability varies widely amongst long term musicians; while some individuals with AP are highly and consistently accurate at pitch labeling, others are less so, despite still performing significantly above chance levels. Additionally, certain musical qualities such as timbre, range, and color (black vs white keys) have been shown to have an effect on AP performance (Bahr, Christensen, & Bahr, 2005; Brammer, 1951; Marvin & Brinkman, 2000; Miyazaki, 1988, 1989, 1990; Takeuchi and Hulse, 1993; Vanzella & Schellenberg, 2010; Wong & Wong, 2014), indicating that AP is not as simple as converting raw frequencies into note names and that these musical qualities may provide contextual cues that can facilitate pitch labeling.

While many studies have assessed the effects of musical qualities such as timbre, range and note color on AP performance, few have incorporated the role of instrument expertise and investigated how it interacts with these qualities to further influence AP performance. The study in Chapter 3 specifically focuses on timbre (piano tones and violin tones) and instrument expertise (pianists and violinists) and investigates how they interact to potentially affect AP performance. Findings indicated that AP performance was indeed impacted as a result of instrument expertise and note timbre. Specifically, AP possessors performed better on trials where the note timbre was congruent with their instrument of expertise, and worse on trials where the note timbre was incongruent with their instrument of expertise. The findings from this study are significant due to the fact that many research studies that test for AP do not account for effects of timbre or instrumental expertise when assessing their participants. This suggests a potential lack of accurate characterization of AP ability in the existing literature, as many individuals with AP may not be performing to the best of their abilities when assessed and as a result, the prevalence of AP may be underreported.

Although the groups described in this dissertation are unrelated, they share the quality of having consistent, highly specific associations with and responses towards particular sounds. By showing how information associated with sounds can influence how we process and respond to those sounds in two separate and rare groups of individuals, the studies in this dissertation make a multidimensional contribution to our understanding of how humans process auditory stimuli.

References

Bahr, N., Christensen, C. A., & Bahr, M. (2005). Diversity of accuracy profiles for absolute pitch recognition. *Psychology of Music*, 33(1), 58–93. doi:10.1177/0305735605048014

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, e10.

Brammer, L. M. (1951). Sensory cues in pitch judgment. *Journal of experimental psychology*, 41(5), 336.

Bruxner, G. (2016). ‘Mastication rage’: a review of misophonia–an under-recognised symptom of psychiatric relevance?. *Australasian Psychiatry*, 24(2), 195-197.

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

Jastreboff, M. M., & Jastreboff, P. J. (2001). Components of decreased sound tolerance: hyperacusis, misophonia, phonophobia. *ITHS News Lett*. 2, 5–7.

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

McGuire, J. F., Wu, M. S., & Storch, E. A. (2015). Cognitive-behavioral therapy for 2 youths with misophonia. *The Journal of Clinical Psychiatry*, 76(5), 573–574.

Marvin, E. W., & Brinkman, A. R. (2000). The effect of key color and timbre on absolute pitch recognition in musical contexts. *Music Perception: An Interdisciplinary Journal*, 18(2), 111-137.

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics*, 44(6), 501–12.

Miyazaki, K. (1989). Absolute Pitch Identification: Effects of Timbre and Pitch Region. *Music Perception*, 7(1), 1–14.

Miyazaki, K. (1990). The speed of musical pitch identification by absolute-pitch possessors. *Music Perception*, 8(2), 177-188.

Potgieter, I., MacDonald, C., Partridge, L., Cima, R., Sheldrake, J., & Hoare, D. J. (2019). Misophonia: A scoping review of research. *J. Clin. Psychol*. doi: 10.1002/jclp.22771.

Rouw, R., & Erfanian, M. (2017). A large-scale study of misophonia. *J. Clin. Psychol*. doi: 10.1002/jclp.22500.

Schröder, A. E., Vulink, N. C., & Denys, D. A. (2013). Misophonia: diagnostic criteria for a new psychiatric disorder. *PLoS ONE* 8:e54706. doi: 10.1371/journal.pone.0054706

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, 113(2), 345–61.

Vanzella, P., & Schellenberg, E. G. (2010). Absolute pitch: Effects of timbre on note-naming ability. *PloS ONE*, 5(11), e15449.

Wong, Y. K., & Wong, A. C. N. (2014). Absolute pitch memory: Its prevalence among musicians and dependence on the testing context. *Psychonomic bulletin & review*, 21(2), 534-542.

Wu, M. S., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2014). Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. *J. Clin. Psychol*. 70, 994–1007. doi: 10.1002/jclp.22098

CHAPTER 1

Misophonia: physiological investigations and case descriptions

Edelstein, M., Brang, D., Rouw, R., and Ramachandran, V.S.

As it appears in

Frontiers in Human Neuroscience

2013

Volume 7, Article 296

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 25 June 2013
doi: 10.3389/fnhum.2013.00296

Misophonia: physiological investigations and case descriptions

Miren Edelstein[1]*, David Brang[1,2], Romke Rouw[3] and Vilayanur S. Ramachandran[1]

[1] Department of Psychology, Center for Brain and Cognition, University of California, San Diego, San Diego, CA, USA
[2] Department of Psychology, Northwestern University, Evanston, IL, USA
[3] Department of Psychology, University of Amsterdam, Amsterdam, Netherlands

Edited by:
Peter G. Enticott, Monash University, Australia

Reviewed by:
Kate Hoy, Monash University, Australia
Arjan Schröder, Academic Medical Center, Netherlands

***Correspondence:**
Miren Edelstein, Department of Psychology, University of California, San Diego, 9500 Gilman Dr. #0109, La Jolla, CA 92093-0109, USA
e-mail: mhedelstein@ucsd.edu

Misophonia is a relatively unexplored chronic condition in which a person experiences autonomic arousal (analogous to an involuntary "fight-or-flight" response) to certain innocuous or repetitive sounds such as chewing, pen clicking, and lip smacking. Misophonics report anxiety, panic, and rage when exposed to trigger sounds, compromising their ability to complete everyday tasks and engage in healthy and normal social interactions. Across two experiments, we measured behavioral and physiological characteristics of the condition. Interviews (Experiment 1) with misophonics showed that the most problematic sounds are generally related to other people's behavior (pen clicking, chewing sounds). Misophonics are however not bothered when they produce these "trigger" sounds themselves, and some report mimicry as a coping strategy. Next, (Experiment 2) we tested the hypothesis that misophonics' subjective experiences evoke an anomalous physiological response to certain auditory stimuli. Misophonic individuals showed heightened ratings and skin conductance responses (SCRs) to auditory, but not visual stimuli, relative to a group of typically developed controls, supporting this general viewpoint and indicating that misophonia is a disorder that produces distinct autonomic effects not seen in typically developed individuals.

Keywords: misophonia, sound sensitivity, skin conductance response, auditory processing, aversive sounds, case reports, autonomic response

GENERAL INTRODUCTION

Misophonia, literally translated to "hatred of sound," is a chronic condition in which specific sounds provoke intense emotional experiences and autonomic arousal within an individual. Trigger stimuli include repetitive and social sounds typically produced by another individual, including chewing, pen clicking, tapping, and lip smacking. These experiences are not merely associative in nature, but drive the sufferer to avoid situations in which they may be produced, limiting one's ability to interact with others and often leading to severe problems in their social and professional lives. Also known as selective sound sensitivity syndrome, the term "misophonia" was first coined by Jastreboff (Jastreboff, 2000; Jastreboff and Jastreboff, 2001a,b, 2003) and little remains known about the condition. To our knowledge only two case studies (Hadjipavlou et al., 2008; Schwartz et al., 2011) and one clinical study (Schröder et al., 2013) have examined misophonia. In the latter study, psychiatrists presented questionnaires and administered interviews to 42 misophonics, an essential first step in showing that misophonia is a primary disorder with no obvious comorbidity with other known psychological or neurological conditions (Schröder et al., 2013).

The prevalence of misophonia is under active investigation but there exist several online support groups with thousands of members (Misophonia UK, Facebook and Yahoo). Sufferers of misophonia are fully aware of its presence and the abnormal responses they have to their trigger sounds. In addition, many sufferers have identified the condition in at least one close relative, suggesting a possible hereditary component. While effective treatments for misophonia remain elusive, individuals report utilizing coping mechanisms to minimize their exposure and response to triggering stimuli (discussed at length below). Further, misophonia appears to exhibit some general similarities to tinnitus. Jastreboff and Hazell (2004) propose that misophonia and tinnitus are both associated with hyperconnectivity between the auditory and limbic systems, suggesting that both conditions would evoke heightened reactions to their respective sounds. However, despite these general similarities, misophonia differs from tinnitus considerably, particularly in terms of how the condition is localized around certain human-produced sounds and situations as opposed to internally perceived, abstract sounds.

While the majority of typically developing individuals experience general and unelaborated emotional reactions to a range of sounds (Halpern et al., 1986), these widespread negative associations remain non-debilitating and at most an annoyance to the listener. One critical possibility is that the valenced associations present in typically developing individuals are matched to those with misophonia, with the latter merely experiencing a more extreme physiological response. Indeed, the sound of fingernails on a chalkboard is an emotionally evocative stimulus that elicits extreme discomfort in the typical population (Zald and Pardo, 2002; Kumar et al., 2012) and misophonic individuals often reference this stimulus in illustrating the extreme nature of their trigger sensations. In this study, we further elaborate on the symptoms and behaviors associated with misophonia as well as examine whether misophonics' physiological responses support their subjective reports of feeling autonomic arousal in response to certain sounds.

EXPERIMENT 1

INTRODUCTION

We first received information about misophonia in December of 2011 through members of an online misophonia support group. From initial descriptions, the condition appeared to have many intriguing qualities in addition to being quite unknown and unexplored. Misophonic individuals were invited to the lab for preliminary interviews with the hope of gaining a more concrete understanding of their experiences with the condition.

MATERIALS AND METHODS

Participants

Eleven individuals with misophonia from the San Diego and Los Angeles areas were recruited from the University of California, San Diego campus, through self-identified contact of our lab as well as through an online misophonia support group (4 males and 7 females, mean age = 35.82; range = 19–65).

Procedure

Thirty to sixty minute semi-structured interviews were conducted by members of our research group on the University of California, San Diego campus. As no set diagnostic criteria for misophonia exists for misophonia, eligibility for study inclusion was based on severity of symptoms paired with experiential descriptions reported by the subject. The five initial interviews were exploratory in nature and included a range of topics, including approximate age of onset, lists of sounds that elicit varying degrees of discomfort, whether or not certain individuals exacerbate the condition, coping mechanisms, common thoughts when experiencing symptoms, physical responses to the trigger sounds, effect of the condition on their daily lives, and other potentially comorbid medical conditions. From these interviews we were able to generate a core set of questions to create the general framework of the subsequent six interviews that were held.

RESULTS AND DISCUSSION

After conducting all 11 interviews, it was apparent that the experiences of the misophonics, though intrinsically variable between subjects, contained noticeable trends and similarities. The most salient categories of assessment and their traits are documented in **Table 1**. In addition, it should be noted that all diagnostic criteria listed by Schröder et al. (2013) were present in the reports of our misophonic subjects (see **Table 1**) even though these interviews were conducted prior to the publishing of that article.

The most important criterion in misophonia is that particular sounds will evoke a disproportional aversive reaction. Our subjects were recruited based on their reports of this characteristic. In accordance with previous reports, our misophonics reported that the worst trigger sounds are chewing, eating, and crunching sounds, followed by lip smacking, pen clicking, and

Table 1 | Summary of qualitative data gathered from interviews of the 11 misophonic subjects (4 males and 7 females, mean age = 35.82; range = 19–65) in Experiment 1, broken down into 18 of the most salient diagnostic categories.

Age of onset	8–10 years old (3)–27% As long as can remember (3)–27% Childhood (3)–27% 17 (1)–9% Early teenage years (1)–9%
Worst trigger sounds	Eating/chewing/crunching sounds (11) Lip smacking (2) Pen clicking (2) Clock ticking (2)
Other trigger sounds	Low frequency bass sounds (8) Pen clicking (4) Footsteps (3) Finger tapping (3) Whistling sounds (3) Typing (3) Lip smacking (2) Clock ticking (1) Plastic bags (1) Repetitive barking (1) Finger tapping (1) Sniffling (1)
Localized around certain individuals?	Yes (9)–82% No (2)–18%
Worsened over time?	Yes (5)–45% Stays the same (3)–27% No, gotten better (2)–18% N/A (1)–9%
Own trigger sounds ok?	Yes (10)–91% Avoids producing own trigger sounds (1)–9%
Repetitive sounds worse	Yes (9)–82% N/A (2)–18%
Runs in family?	Yes (6)–55% Not known (3)–27% N/A (2)–18%
Coping strategies	Avoiding or removing self from certain situations (7) (*D,E) Mimicry to "cancel out" sound or retaliate (6) Earplugs/headsets/music (6) Is conscientious about own sounds (5) Distract self (5) Ask others to stop (4) Positive internal dialog (1)
Effect of alcohol/caffeine	Alcohol lessens symptoms (7) Caffeine worsens symptoms (4) Symptoms not affected by caffeine (2) Does not use caffeine (2) Does not use alcohol (2) Symptoms not affected by alcohol (1) N/A (1)

(Continued)

Table 1 | Continued

Physical locations and descriptions of discomfort (*A)	Pressure in chest, arms, head, or whole body (5) Clenched/tightened/tense muscles (5) Increase in body temperature, blood pressure, or heart rate (2) Pained by trigger sounds (1) Hard to breathe (1) Sweaty palms (1)
Visual triggers	Jiggling/swinging legs (5)
Bothered by Ss sounds	Yes (6)–55% N/A (3)–27% No (2)–18%
Feelings and emotions associated with trigger sounds*	Sounds are invasive, intrusive, insulting, violating, offensive, disgusting, rude (9) (*A,D) Stress/anxiety (5) Anger or rage (4) (*D) Extreme annoyance/irritation (4) (*A,D) Panic (2) (*B) Impatience (1) Aggravation (1) (*D) Feeling trapped (1) (*B)
Other potentially comorbid medical conditions (*F)	Tinnitus (2) Obsessive-compulsive personality traits (2) Hyperacusis (1) Auditory processing disorder (1) ADD (1) PTSD (1) None (6)
Bothered by sounds produced by animals or children	Yes (1)–9% No (8)–73% N/A (2)–18%
Thoughts when experiencing trigger sounds	"I want to punch this person" "I hate this person" "Why won't they stop? I don't want to hurt their feelings by changing seats" (*C) "Why are they eating that way?" "Why are you doing that? It's rude" "Would you shut up?" "Stop it, I can't stand it" "Don't you know what you sound like?" "Why am I like this?" (*C) "Are they doing this on purpose?" "Why does he have to _____ so loudly?" "They should be more conscious of how they're affecting others" "I envy people who aren't bothered by sounds" (*C)
Effect on life	Realizes they are hyper focused on noises that should be in the background and are unable to ignore them (9) (*C,E) Cannot pay attention at a movie or in class when people are making trigger sounds (8) (*E)

(Continued)

Table 1 | Continued

	Tries not to be around people if they make trigger sounds (7) (*D,E) Can be triggered by sounds from television or videos (7) (*E) Triggers are worse when tired (7) Stays away from certain foods/avoids making certain sounds (3) (*D,E) Feels better when can locate source of sound (3) Thoughts of suicide (1)

The number of subjects reporting a criterion can be found in parentheses to the right of each description.
Criteria marked with an asterisk (*) designate diagnostic criteria (A–F) consistent with those proposed by Schröder et al. (2013).
Please see General Discussion for more details.

clock ticking (see **Table 1**). Other notable trigger sounds include low frequency bass sounds, footsteps, finger tapping, whistling sounds, and typing (see **Table 1**). Nine of our 11 misophonics reported that sounds repetitive in nature were particularly bad. In addition, six of our misophonics indicated that spoken "Ss" sounds were unpleasant, although not quite on the same level as trigger sounds.

In terms of aversive responses to these sounds, misophonics report a range of negative feelings, thoughts, as well as physical reactions. Some of the negative feelings experienced include intense anxiety, panic, anger, extreme irritation, and even rage (see **Table 1**). Additionally, in the context of our study, it is important to distinguish anxiety from fear. Specifically, while our subjects report feeling extreme stress and anxiety in response to trigger sounds, they did not report being afraid or fearful of them. Nine of our 11 misophonics reported trigger sounds as being invasive, intrusive, disgusting, or rude. They also reported feeling offended or violated by these sounds to the point where negative thoughts such as "I hate this person," "Stop it, I can't stand it," and "Don't you know what you sound like?" enter their minds. However, 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. In line with this, the worst situations for misophonics are often ones where they feel trapped and unable to escape, including long trips in cars or planes. Similarly, two misophonics report that trigger sounds at school or at home are worse than in places one can easily leave, such as a public place. However, despite extreme discomfort, misophonics generally do not physically act out on feelings of aggression. Some report instances of snapping at others while others internalize their frustration (see **Table 1**).

A final indication that misophonia produces physical and autonomic responses is the suggestion that pharmacological agents affect the condition. Four of our misophonics indicated that caffeine intensifies misophonic experiences while seven of misophonic individuals indicated that alcohol decreases symptomatology; these subjects describe that while under the influence of alcohol they can still hear the sound but their aversive response is not as strong.

In response to their aversive reactions to trigger sounds, misophonic individuals have developed a number of coping strategies including: avoiding or removing themselves from certain situations, mimicking trigger sounds, or the action producing it to "cancel out" or "retaliate," utilizing earplugs, headsets or listening to music, distracting oneself, reciting positive internal dialog to help calm themselves, asking others to stop making the sounds, as well as being conscientious about their own sounds (see **Table 1**).

The degree to which quality of life is affected varied between our misophonic participants. One subject reported that misophonia "…does not affect the quality of my life too much. But it seems ridiculous and I would like to get rid of it" while another subject reported that misophonia had in the past evoked thoughts of suicide. These reports indicate there might be different degrees of the misophonic condition, ranging from mildly hindering to severely debilitating.

Misophonic individuals most commonly describe onset of the condition in childhood. Two subjects reported that with age, they learned to better cope with their misophonia, five subjects reported that it worsened over time (due to increasing aversiveness as well as increasing number of triggering stimuli) and three recalled no change over time. It is not fully understood why differences in trigger accumulation and severity develop between misophonics but it appears that prolonged and repeated exposure to a sound may be a contributing factor. For example, one of our misophonic subjects related this to the "honeymoon" period in a new job or relationship, in which for a few years new sounds caused little irritation. However, over time the negative affect of these sounds intensified to become triggers as well.

Six of our misophonics reported that one or several close family members display misophonic-like symptoms and behaviors. Two subjects had no information on this topic and three reported that they do not believe that misophonia runs in their families. While these reports are only anecdotal, they suggest there may be a familial or genetic component to misophonia, calling for further investigation in future studies.

Interestingly, misophonic individuals further report that responses evoked by trigger sounds appear to be modulated by prior knowledge, context, and sound source, implying that the condition is not driven simply by the physical properties of sound alone. For example, nine of our misophonics indicated that their misophonia is isolated to or exacerbated by certain individuals, usually close friends, coworkers, or family members whom they are exposed to frequently (see **Table 1**). Another curious characteristic described by 10 of our misophonics is the fact that self-induced trigger sounds (trigger sounds produced by the misophonic individual themselves) will not evoke nearly as much of an aversive response as when produced by others. In fact, as mentioned earlier, mimicking trigger sounds is one of the coping strategies utilized by misophonics to "overwrite" the disturbing sound being produced by another individual. Several misophonics even report eating foods in synchrony with the other person. However, mimicking is also mentioned as a way to retaliate against the offending individual producing the sounds, thus acting as a way to cope with the anger evoked by the condition.

The interviews further revealed an interesting effect of the role of context on aversive responses. For instance, eight of our misophonics report eating and chewing sounds (severely offensive triggers associated with rudeness when produced by human adults) will not bother them nearly as much if produced by animals or babies (see **Table 1**). One individual described that, as these individuals have little control over their actions and "don't know any better," it helps in cancelling out strong aversive feelings. These results suggest that the aversive responses experienced by misophonics are explicitly tied to other individuals, implying an underlying social component to the condition. Accordingly, even though our subjects fit in with Schröder et al.'s (2013) diagnostic criterion of misophonics being aware of their condition, and recognizing their feelings as "excessive, unreasonable, or out of proportion," they will still comment on the inappropriateness of another person's behavior nonetheless.

Another recurring topic from the interviews is the role of attention in misophonia. Nine of our misophonics report being hyper-focused on sounds that normally exist as background noise. One misophonic subject described the inability to tune out background noises as being like an "involuntary cocktail party effect" while another mentioned that "noises are never in the background. People sounds crash right through jet engine sounds." Eight of our misophonics described being unable to pay attention to a movie or lecture when individuals around them produce trigger sounds, with partial remediation by distracting themselves and directing their attention elsewhere. In addition, it is possible that through understanding the role of attention in misophonia, potential treatments may be able to be developed.

In accordance with Schröder et al. (2013), our subjects reported a few symptoms shared with other diagnoses, however the complete symptomology of misophonia does not fit with any of the diagnostic categories in the diagnostic and statistical manual of mental disorders (DSM-IV). In their interviews, subjects described symptoms related to obsessive-compulsive disorder (OCD), attention deficit disorder (ADD), post-traumatic stress disorder (PTSD), auditory processing disorders as well as tinnitus and hyperacusis (see **Table 1**). However, these symptoms did not cover the full range of complaints, including the critical symptom of misophonia (a strong aversive response to particular sounds). Two of our misophonics reported being treated with medications, including antianxiety medications and antidepressants, that were intended to alleviate some of the effects of misophonia but as it stands, a treatment to fully address the root of the problem still remains elusive. Thus, our results are in line with the previous conclusion that misophonia is not part of another clinical, psychiatric, or psychological disorder (Schröder et al., 2013).

EXPERIMENT 2

INTRODUCTION

Qualitative assessments of misophonic subjects demonstrated the consistent association between specific sounds and intense emotional experiences. In order to confirm the presence of these emotional reactions and further examine their relationship to sound preferences present in the general population, we measured skin conductance response (SCR) while misophonic participants and typically developed individuals were exposed to aversive and non-aversive auditory, visual, and auditory-visual stimuli. SCR measures the electrical conductance of the skin and consequently the amount of sweat produced. Because sweat production is not under volitional control, SCR is widely accepted to indicate arousal of the sympathetic nervous system (Critcheley, 2002). For these reasons, we believe SCR to be an appropriate method of measuring autonomic arousal to various emotion-eliciting stimuli.

MATERIALS AND METHODS

Participants

Six misophonic subjects who also participated in Experiment 1 (2 males and 4 females; mean age = 22.8; range = 19–30) and five controls (mean age = 22; range = 19–29) matched on age and gender participated in the experiment; A sixth control was excluded due to an error during data collection. Controls were recruited from the student population at the University of California, San Diego. All participants reported normal hearing and vision, gave signed, informed consent prior to the experiment, and participated either for cash or in fulfillment of a course requirement. The study was reviewed and approved by the university's Human Research Protections Program. Total experiment time was less than 1 h.

Procedure and stimuli

Participants were seated 20 inches from an 18 inch monitor and provided Sennheiser® headsets. SCR recordings were acquired with BIOPAC System (MP100A-CE) and AcqKnowledge 4.1 recording software. A pair of Ag-AgCl electrodes was attached to the palmar surface of the middle and ring fingers of the participant's dominant hand. Prior to attachment, participants' hands were cleansed with an alcohol wipe and a skin conductance gel was applied to each electrode. SCR was recorded in micro Siemens at a rate of 30 samples/s. Participants were instructed to relax with their dominant hand placed palm up on their thigh and to minimize movement throughout the duration of the experiment. SCR was examined in subjects prior to experimental testing for typicality; absence of a normal response precluded a subjects' participation in the rest of the study.

Stimuli included 31 video clips either acquired from YouTube or recorded in the lab. Video content varied in order to cover a range of sounds and predicted emotional responses in misophonic subjects, selected based on interview data from Experiment 1. Example stimuli included birds singing, children laughing, whale song, nails on a chalkboard, lips smacking, gum chewing, etc. Each clip lasted for 15 s. Auditory and visual components of these videos were separated to generate auditory alone, visual alone, and auditory-visual conditions. Each auditory, visual, and auditory-visual stimulus was presented once for a total of 93 trials. Trial order was randomized into two orders and order was counterbalanced across participants. Critically, as each specific video was presented a total of three times (once in each auditory, visual, and auditory-visual condition), a consistent ordering of the presentation of each stimulus was maintained for each type: auditory alone, visual alone, followed by auditory-visual. Stimuli were presented with E-Prime® version 2.0.

On each trial, participants viewed a centrally presented fixation cross for a 5-s period, followed by either an auditory clip (A), visual movie (V), or auditory-visual movie (AV) for 15 s, concluded with an inter-trial interval of 10 s; during this 10-s interval subjects provided a verbal aversiveness rating on a scale of 0–4 based on how much discomfort they experienced in response to the preceding trial. Participants were informed that a rating of 0 would signify no discomfort at all and a rating of 4 would signify an extreme amount of discomfort, anxiety, or an urge to leave the room. Each aversiveness rating was recorded by the experimenter.

Data preprocessing

As our stimuli were presented in quick succession, a linear downward trend was observed throughout the recording session. To account for this artifact, separate linear regressions were fitted to the 5-s fixation period at the start of each trial through a line of best fit. Each observed value during the stimulus epoch was re-plotted as the residual of this line of best fit, normalizing for the pre-stimulus baseline period and removing artifact trends present throughout the epoch. A consistent pattern of results was additionally observed on non-detrended data.

Data analysis

SCR onset was time-locked to pre-stimulus fixation cross. Mean SCR was calculated from the 15-s stimulus epoch for each trial, following the fixation cross. Mean values exceeding three standard deviations from the mean SCR across all trials for each participant were deemed outliers and consequently removed from the dataset; an average of 1.9% of trials were removed per participant.

Statistical analyses

First, we conducted repeated measures ANOVAs across factors of Group (misophonics, controls), Measurement (SCR, aversiveness rating), and Condition (auditory, visual, auditory-visual) to observe overall effects. Follow-up ANOVAs, non-parametric independent samples tests and descriptive analyses were conducted to explore group differences. Follow-up correlations revealed further group differences as well as similarities. Greenhouse-Geisser corrections were used where appropriate, but we report the original degrees of freedom for clarity.

RESULTS

Overall group effects

As an overall examination of the data, we conducted a repeated measures ANOVA with factors Group (misophonics, controls), Measurement (SCR, subjective rating), and Condition (auditory, visual, auditory-visual). Results showed significant main effects of Group [$F_{(1, 9)} = 17.5, p < 0.005$], Condition [$F_{(2, 18)} =$

47.3, $p < 0.001$], and Measurement [$F_{(1, 9)} = 48.5$, $p < 0.001$], as well as significant interactions between Group × Condition [$F_{(2, 18)} = 18.8$, $p < 0.005$], Group × Measurement [$F_{(1, 9)} = 13.7$, $p < 0.01$], Measurement × Condition [$F_{(2, 18)} = 40.5$, $p < 0.001$], and Group × Measurement × Condition [$F_{(2, 18)} = 16.2$, $p < 0.005$].

However, as the primary goal of this study was to examine unisensory responses to stimuli in both groups, subsequent tests for group effects excluded multisensory (auditory-visual) trials and included only auditory and visual conditions. **Figure 1** shows misophonic and control subjects' average SCR data in auditory and visual conditions as a function of time. A repeated measures ANOVA with factors of Group (misophonics, controls), Measurement (SCR, subjective rating), and Condition (auditory, visual) similarly identified significant main effects of Group [$F_{(1, 9)} = 14.3$, $p < 0.005$], Condition [$F_{(1, 9)} = 47.5$, $p < 0.001$], and Measurement [$F_{(1, 9)} = 40.7$, $p < 0.001$], as well as significant interactions between Group × Condition [$F_{(1, 9)} = 17.5$, $p < 0.005$], Group × Measurement [$F_{(1, 9)} = 10.1$, $p < 0.05$], Measurement × Condition [$F_{(1, 9)} = 44.0$, $p < 0.001$], and Group × Measurement × Condition [$F_{(1, 9)} = 16.1$, $p < 0.005$]. This overall ANOVA validated the use of follow-up analyses to test specific hypotheses.

Group differences

We conducted additional follow-up repeated measure ANOVAs with factors of Group (misophonics, controls) and Condition (auditory, visual), first for subjective aversiveness ratings alone. Results showed main effects of Group [$F_{(1, 9)} = 12.4$, $p < 0.01$] and Condition [$F_{(1, 9)} = 46.5$, $p < 0.001$], and critically an interaction between the two [$F_{(1, 9)} = 17.1$, $p < 0.005$] supporting the differences between the groups (see **Figure 2A**). This difference between the groups was largely due to controls rarely rating stimuli as greater than 2 on the aversiveness scale (ranging from 0 to 4; see **Figures 3A,B**). Examining this model for SCR data yielded a similar pattern of results with main effects of Group [$F_{(1, 9)} = 6.77$, $p < 0.05$] and Condition [$F_{(1, 9)} = 11.9$, $p < 0.01$], and a marginally significant interaction between the two [$F_{(1, 9)} = 4.53$, $p = 0.06$] (see **Figure 2B**).

Given the small sample size of these groups, follow-up non-parametric independent-samples Mann-Whitney *U*-tests were used to compare groups across these critical conditions. Misophonics reported significantly higher ratings than control subjects in response to auditory stimuli, $U_{(9)} = 29.0$, $p < 0.01$, but not visual stimuli, $U_{(9)} = 23.5$, $p = 0.13$. The median rating of auditory trials was 1.82 ($SD = 1.38$) for misophonics and 0.42 ($SD = 0.77$) for controls while the median rating of visual trials was 0.29 ($SD = 0.98$) for misophonics and 0.19 ($SD = 0.55$) for controls. This pattern of results was consistent with SCR responses, with misophonics producing larger SCR responses than controls to auditory stimuli, $U_{(9)} = 28.0$, $p < 0.05$, but not visual stimuli, $U_{(9)} = 21.0$, $p = 0.33$. The median SCR of auditory trials was 0.15 micro Siemens ($SD = 0.40$) for misophonics and 0.03 micro Siemens ($SD = 0.11$) for controls while the median SCR of visual trials was 0.07 micro Siemens ($SD = 0.39$) for misophonics and 0.00 micro Siemens ($SD = 0.08$) for controls. The same pattern of results for these tests was observed with parametric independent samples *t*-tests.

In order to determine if higher SCR is directly correlated with higher aversiveness ratings, we examined individual subjects' aversiveness ratings relative to average SCR activity from all auditory, visual, and auditory-visual trials. Results identified a significant positive correlation between average aversiveness ratings and average SCR across all participants (see **Figure 4**), ($r_s = 0.700$, $N = 11$, $Z = 2.21$, $p < 0.05$), indicating that stimuli subjectively thought of as aversive generally evoked a proportional SCR.

Group similarities

As an examination of whether the stimuli that trigger aversive experiences in misophonic individuals are idiosyncratic to the condition or consistent to, though more extreme than,

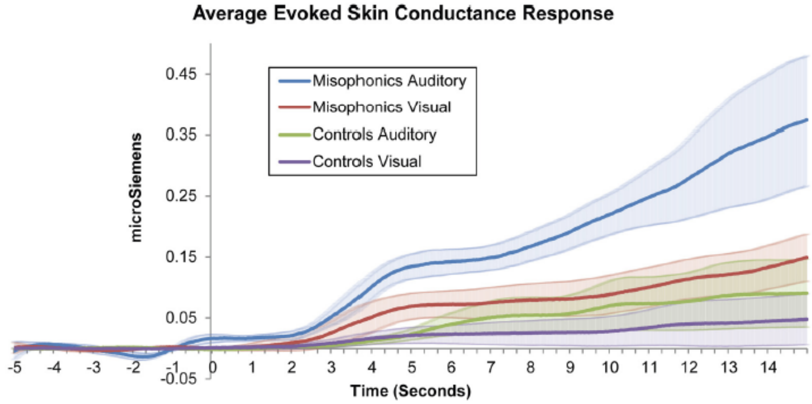

FIGURE 1 | Average misophonic and control participants' skin conductance response to auditory and visual stimuli as a function of time.

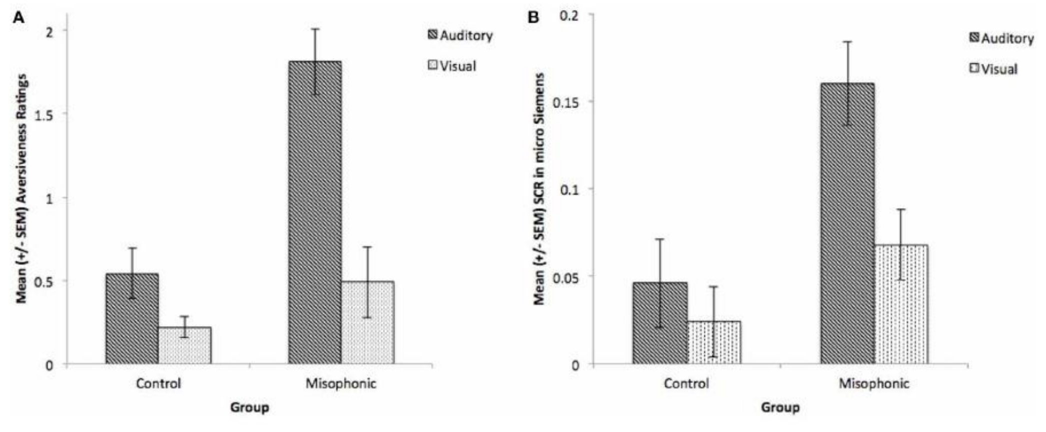

FIGURE 2 | Group means of controls and misophonics, per presentation condition (auditory and visual) for (A) subjective reports and (B) SCR.

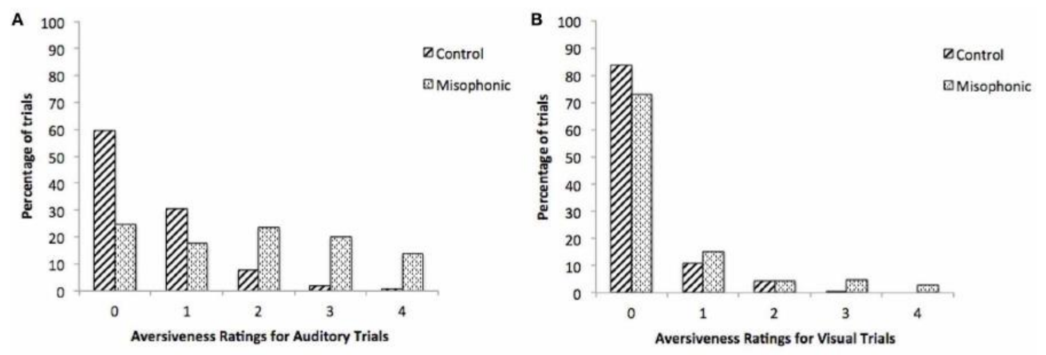

FIGURE 3 | Percentage of trials per index on the 5-point aversiveness scale, for controls and misophonics, during (A) auditory and (B) visual conditions.

preferences present in the general population, we examined the consistency of ratings across the groups. Findings indicated a significant positive correlation between misophonic and control aversiveness ratings across all three types of stimuli, ($r_s = 0.605$, $N = 93$, $Z = 5.80$, $p < 0.001$); this correlation is additionally present when examining the correlation between the groups for only auditory trials, ($r_s = 0.413$, $N = 31$, $Z = 2.26$, $p < 0.05$; see **Figure 5**) suggesting that misophonics and controls find similar stimuli to be aversive and non-aversive.

DISCUSSION

Experiment 2 provides, to the best of our knowledge, the first experimental investigation on misophonia, serving to validate the severity of this chronic condition beyond anecdotal description. Misophonic subjects rated auditory stimuli as more aversive than the same visual stimuli, and this pattern was consistent with SCR measurements. Furthermore, SCR and subjective ratings to auditory stimuli were greater in misophonic individuals than controls, supporting the specificity of aversive reactions in misophonia. Nevertheless, misophonic subjects demonstrated increased ratings and SCR regardless of stimulus type, as revealed by observed main effects of group, possibly denoting generalized anxiety to the stimuli used in the present study.

The significant positive correlation between average aversiveness ratings and mean SCR across all participants importantly confirms the validity of each subject's ratings during the task. Therefore, participant's physiological responses to stimuli were consistent with their subjective ratings. However, as shown in **Figure 4**, this positive correlation seems most likely driven by group differences between misophonics, (represented in green) and controls (represented in blue).

The significant positive correlation between misophonic aversiveness ratings and control aversiveness ratings reflects a general agreement of the relative valence of the inducing stimuli across the groups. In other words, misophonics and controls find similar stimuli to be aversive and non-aversive on a subjective level, suggesting that misophonics may experience an extreme form of the discomfort most individuals experience to normally aversive or irritating stimuli. This raises the important possibility that there is nothing intrinsically different about misophonic individuals from those in the general population and misophonic individuals are merely at the tail end of the distribution.

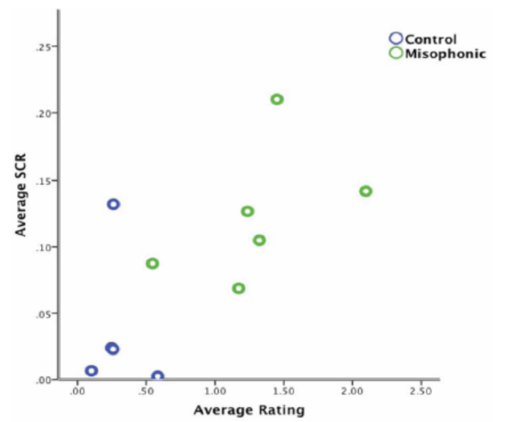

FIGURE 4 | Correlation of average aversiveness ratings and average SCR (in micro Siemens) for all trials across all subjects.

GENERAL DISCUSSION

In a preliminary examination of individuals with misophonia, we report qualitative and physiological investigations of the condition and its relationship to responses in the typical population. Experiment 1, which is comprised of qualitative assessments on eleven misophonic subjects, examined the qualities associated with misophonia in order to help develop reliable diagnostic criteria and understand the complex social factors involved. Results were consistent with early reports of the phenomenon, such as the critical characteristic of misophonia being a disproportionately aversive reaction is in response to common sounds in everyday life. Additionally, a visceral autonomic response is physically felt in misophonics in response to trigger sounds. In Experiment 2, physiological measurements were acquired on six misophonic individuals using SCR to provide an objective corroboration of misophonics' reports that specific sounds evoke intense emotional reactions. Results showed an increased autonomic response to trigger sounds, but not visual stimuli, in misophonics as compared with non-misophonic controls.

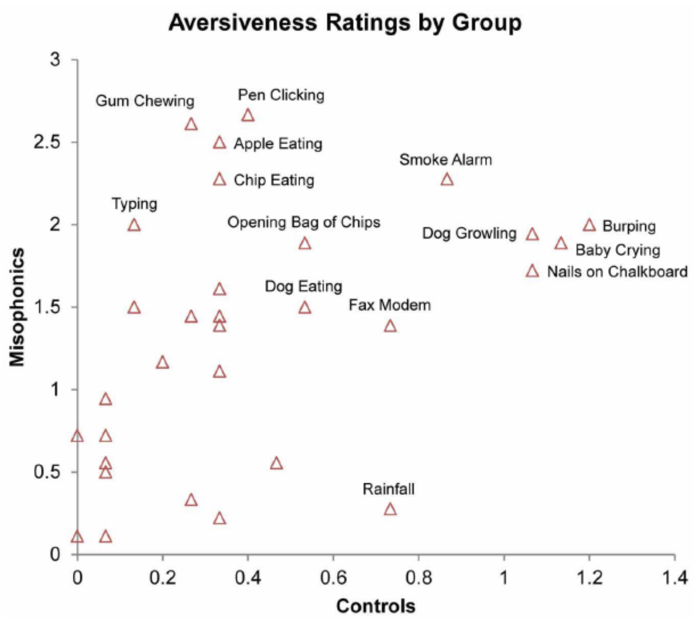

FIGURE 5 | Correlation of average aversiveness ratings of stimuli (x- and y-axis) across conditions in misophonics and controls. Select stimuli identified by proximal text.

Administering semi-structured interviews proved to be an effective way of determining the most critical symptoms, triggers and experiences associated with misophonia as well as the degree to which these varied across subjects. In addition to reporting psychological symptoms, all of our misophonics reported physical symptoms synonymous with autonomic arousal in response to trigger sounds. Furthermore, our qualitative results are in line with all of the diagnostic criteria proposed by Schröder et al. (2013) which, shortly summarized are: (A) aversive and angry feelings evoked by particular sounds, (B) rare potentially aggressive outbursts, (C) recognition by the misophonic individual that his/her behavior is excessive, (D) avoidance behavior, (E) distress and interference in daily life, and lastly, (F) the lack of another condition to account for all symptoms. Additionally, our principal finding that misophonic individuals experience physical, autonomic arousal that is measurable by SCR, provides empirical validation for some of the aforementioned critical criteria proposed by Schröder et al. (2013), particularly criterion A. Through conducting interviews, we also identified other interesting aspects of misophonia that were not previously apparent. In particular, subjects reported that misophonia can be modulated by social expectations as well as situational context, indicating that the condition may be more complicated than merely an aversive response to the purely physical properties of sounds. Additionally, the finding that misophonic individuals report involuntary, physiological distress in response to a very specific subset of social sounds supplements research on complex mind-body interactions, with high-level knowledge demonstrating prolonged and specific physiological reactions (e.g., as in placebos; Margo, 1999). However, at this time, these speculations remain based on anecdotes and need to be properly tested in the future before firm conclusions can be drawn.

To date, no research has examined the neurological origin of misophonia, and preliminary investigations suggest it is not due to any primary neurological or psychological disorder or trauma (Schröder et al., 2013). Nevertheless, misophonia displays similarities to a genetic condition known as synesthesia. In synesthesia, as in misophonia, particular sensory stimuli evoke particular and consistent, additional sensations and associations. Well-known forms of synesthesia include letters evoking a particular color, or sounds/music evoking colors (Cytowic, 1989; Baron-Cohen et al., 1996; Simner et al., 2006) but there are in fact many different subtypes of synesthesia, with a variety of "inducers" (e.g., music, taste, words, sequences) evoking certain "concurrents" (e.g., color, shapes, taste). While most synesthesia research has examined the perceptual sensations related to synesthesia, the condition seems to have an affective component as well. First, synesthetic congruency (e.g., when a grapheme-color synesthete sees a letter in the "correct" color) is related to positive affect (e.g., Callejas et al., 2007). Furthermore, both inducers (Ward, 2004; Ramachandran et al., 2012) and concurrents (Simner and Holenstein, 2007) can be of emotional rather than perceptual nature. Interestingly, the latter indicates that for certain subtypes of synesthesia, similar to misophonia, inducers evoke a particular feeling or emotion rather than a pure perceptual sensation. This has been studied in tactile-emotion synesthesia (e.g., feeling sandpaper evokes a feeling of jealousy; Ramachandran and Brang, 2008). Synesthetic associations, like misophonic experiences, are automatic (in the sense that they do not take effort or conscious deliberation), are consistent within an individual and persist throughout life, and seem to run in families (Asher et al., 2009; Tomson et al., 2011; for a review see Brang and Ramachandran, 2011). Given these similarities, neuroimaging findings in synesthetes may provide us with hypotheses on the neural basis of misophonia. First, associated sensations in synesthesia are found to be associated with co-activation in relevant (associated) brain areas (Nunn et al., 2002; Hubbard et al., 2005; Rouw et al., 2011). Furthermore, previous studies support a direct linking of relevant sensory regions in synesthesia (Hubbard and Ramachandran, 2001), mediated by an actual increase of anatomical connectivity (Rouw and Scholte, 2007; Zamm et al., 2013). Similarly, altered connections from a lesioned thalamus to the cerebral cortex (Ro et al., 2007; Beauchamp and Ro, 2008) led to a type of acquired synesthesia in which auditory stimuli produced tactile percepts. Differing in the level of specificity and complexity of evoked responses observed in synesthetes, individuals with misophonia display basic and non-elaborated responses to triggering stimuli, varying largely in the intensity of the response. Nevertheless, the underlying neurological cause of this condition may be similar to that of synesthesia in terms of enhanced connectivity between relevant brain regions. In short, a pathological distortion of connections between the auditory cortex and limbic structures could cause a form of sound-emotion synesthesia.

This study also provides the critical finding of a relationship between aversive stimuli in misophonia and mildly aversive stimuli in the general population. That is, in Experiment 2 we observed a significant correlation between aversive ratings across the groups, suggesting that misophonia may be based on mechanisms fundamentally present in the general population, but simply exaggerated in misophonia. Critically, as observed in the interviews in Experiment 1, many of the common aversive stimuli in misophonia are also deemed as socially inappropriate in western society (e.g., lip smacking, repetitive tapping, etc.). While speculative at present, this consistent pattern raises the possibility that the aversive nature of these stimuli to all individuals may be based on the same driving factors (though notably more mild) as in misophonia, leading to the development of these cultural norms.

The present paradigm was designed to include a range of aversive stimuli for misophonic individuals based on our preliminary interviews in Experiment 1. Accordingly, misophonic individuals reported a large number of the stimuli as aversive: mean 24.2% and median 24.7% stimuli with a rating of 3 or 4. In contrast, control participants reported very few stimuli as very aversive: mean 2.4% and median 0.0% stimuli with a rating of 3 or 4 (**Figures 3A,B**). Potential future studies are suggested to examine if this same pattern of group differences is consistent with stimuli that evoke a broader range of aversive responses in typically developed individuals.

As the current study is exploratory in nature and included a small sample of participants, there are several limitations to acknowledge. One limitation is that the presentation of stimuli in a controlled laboratory setting lacked the ecological validity of how these stimuli occur in the real world. As such, several misophonics reported that because they knew each clip would end in a matter of seconds, their physiological reactions were tempered, consistent with self-reports in Experiment 1 showing that contextual information about these cues mediated subjects' responses. We predict naturalistic observational studies of physiological reactions in misophonic individuals will show a similar but more extreme pattern of results to those observed here. A second limitation is that while SCR is a good measure of autonomic arousal in response to emotion-eliciting stimuli, it does not indicate what specific emotion is being experienced at the time. Instead it only indicates a very general, physiological arousal that can be interpreted in many ways. For example, SCR would not be able to differentiate anxiety and aggression. However, information as to what exactly a subject was feeling during each stimulus can potentially be inferred by obtained self-reports after each trial. A third limitation is the fact that no rigorous diagnostic tests or screenings were utilized during interviews to completely exclude the possibility that subjects' symptoms were being driven by another underlying condition. Also, interviews were conducted by members of our research group and not by psychiatrists. Potentially comorbid conditions were therefore determined from the self-reports of subjects (some of whom had previous, official diagnoses), and the discretion of the researchers. However, because these interviews were not conducted with the intent of being clinical or diagnostic in nature, but rather to gain more insight into the phenomenological experiences of individuals who identify with having misophonia, we believe these findings are still of considerable value to the research community and misophonic individuals alike. A fourth limitation of the study is the small sample size. As research on misophonia is limited to the last few years and little remains known about the condition, obtaining a large sample size for this study was not feasible. Nevertheless, while these results should be validated on a larger group of subjects, we believe they reflect properties of the condition generalizable to the misophonia community in general.

While these data serve to support the veracity of the subjective reports in misophonia as an intrusive and labile condition, numerous additional avenues remain for future research. Critically, as this condition appears to be chronic, the nature of how subjects' triggers evolve over time should be investigated. How does context contribute to and modulate misophonia and can contextual information or expectation effects bias subjects' responses to aversive stimuli? Critically, what are the mechanisms (genetic, neurological, and/or psychological) that underlie the condition? While speculative at present, one potential neural mechanism for misophonia may lie in aberrant anatomical or functional connections between auditory and limbic regions, akin to the finding of increased structural connectivity in synesthesia. Regardless of the mechanisms that underlie misophonia, the present research supports its validity as an intrusive condition and highlights the need for additional research into contributing factors and potential treatments.

ACKNOWLEDGMENTS

The authors thank Laura Case for contributing to the experimental design, Dr. Julian Parris for assisting with data analysis, Ronald E. Robertson and Marli Gitelson for assisting in subject recruitment and data collection, and all of our subjects who kindly agreed to participate.

REFERENCES

Asher, J. E., Lamb, J. A., Brocklebank, D., Cazier, J. B., Maestrini, E., Addis, L., et al. (2009). A whole-genome scan and fine-mapping linkage study of auditory-visual synesthesia reveals evidence of linkage to chromosomes 2q24, 5q33, 6p12, and 12p12. *Am. J. Hum. Genet.* 84, 279–285. doi: 10.1016/j.ajhg.2009.01.012

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., and Bolton, P. (1996). Synaesthesia: prevalence and familiarity. *Perception* 25, 1073–1079. doi: 10.1068/p251073

Beauchamp, M. S., and Ro, T. (2008). Neural substrates of sound-touch synesthesia after a thalamic lesion. *J. Neurosci.* 28, 13696–13702. doi: 10.1523/JNEUROSCI.3872-08.2008

Brang, D., and Ramachandran, V. S. (2011). Survival of the synesthesia gene: why do people hear colors and taste words. *PLoS Biol.* 9:e1001205. doi: 10.1371/journal.pbio.1001205

Callejas, A., Acosta, A., and Lupiáñez, J. (2007). Green love is ugly: emotions elicited by synestheticgrapheme-color perceptions. *Brain Res.* 1127, 99–107. doi: 10.1016/j.brainres.2006.10.013

Critcheley, H. D. (2002). Electrodermal responses: what happens in the brain. *Neuroscientist* 8, 132–142. doi: 10.1177/107385840200800209

Cytowic, R. E. (1989). *Synaesthesia: A Union of the Senses.* New York, NY: Springer-Verlag.

Hadjipavlou, G., Baer, S., Lau, A., and Howard, A. (2008). Selective sound intolerance and emotional distress: what every clinician should hear. *Psychosom. Med.* 70, 739–740. doi: 10.1097/PSY.0b013e318180edc2

Halpern, L., Blake, R., and Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Percept. Psychophys.* 39, 77–80. doi: 10.3758/BF03211488

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., and Boynton, G. M. (2005). Individual differences among grapheme-color synesthetes: brain-behavior correlations. *Neuron* 45, 975–985. doi: 10.1016/j.neuron.2005.02.008

Hubbard, E. M., and Ramachandran, V. S. (2001). Cross wiring and the neural basis of synaesthesia. *Invest. Ophthalmol. Vis. Sci.* 42, S712.

Jastreboff, M. M., and Jastreboff, P. J. (2001a). Components of decreased sound tolerance: hyperacusis, misophonia, phonophobia. *ITHS News Lett.* 2, 5–7.

Jastreboff, M. M., and Jastreboff, P. J. (2001b). Hyperacusis. *Audiology Online.* June 2001. Available online at: http://www.audiologyonline.com

Jastreboff, P. J. (2000). "Tinnitus habituation therapy (THT) and tinnitus retraining therapy (TRT)," in *Tinnitus Handbook*, ed R. S. Tyler (San Diego, CA: Singular, Thomson Learning), 357–376.

Jastreboff, P. J., and Hazell, J. (2004). *Tinnitus Retraining Therapy: Implementing the Neurophysiological Model.* New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511544989

Jastreboff, P. J., and Jastreboff, M. M. (2003). Tinnitus retraining therapy for patients with tinnitus and decreased sound tolerance. *Otolaryngol. Clin. North Am.* 26, 321–336. doi: 10.1016/S0030-6665(02)00172-X

Kumar, S., von Kriegstein, K., Friston, K., and Griffiths, T. D. (2012). Features versus feelings: dissociable representations of the acoustic features and valence of aversive sounds. *J. Neurosci.* 32, 14184–14192. doi: 10.1523/JNEUROSCI.1759-12.2012

Margo, C. E. (1999). The placebo effect. *Surv. Ophthalmol.* 44, 33–44. doi: 10.1016/S0039-6257(99)00060-0

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., et al. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nat. Neurosci.* 5, 371–375. doi: 10.1038/nn818

Ramachandran, V. S., and Brang, D. (2008). Tactile-emotion synesthesia. *Neurocase* 14, 390–399. doi: 10.1080/13554790802363746

Ramachandran, V. S., Miller, L. E., Livingstone, M. S., and Brang, D. (2012). Colored halos around faces and emotion-evoked colors: a new form of synesthesia.

Neurocase 18, 352–358. doi: 10.1080/13554794.2011.608366

Ro, T., Farne, A., Johnson, R. M., Wedeen, V., Chu, Z., Wang, Z. J., et al. (2007). Feeling sounds after a thalamic lesion. *Ann. Neurol.* 62, 433–441. doi: 10.1002/ana.21219

Rouw, R., and Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nat. Neurosci.* 10, 792–797. doi: 10.1038/nn1906

Rouw, R., Scholte, H. S., and Colizoli, O. (2011). Brain areas involved in synaesthesia: a review. *J. Neuropsychol.* 5, 214–242. doi: 10.1111/j.1748-6653.2011.02006.x

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

Schwartz, P., Leyendecker, J., and Conlon, M. (2011). Hyperacusis and misophonia: the lesser-known siblings of tinnitus. *Minn. Med.* 94, 42–43.

Simner, J., and Holenstein, E. (2007), Ordinal linguistic personification as a variant of synesthesia. *J. Cogn. Neurosci.* 19, 694–703. doi: 10.1162/jocn.2007.19.4.694

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., et al. (2006). Synaesthesia: the prevalence of atypical cross-modal experiences. *Perception* 35, 1024–1033. doi: 10.1068/p5469

Tomson, S., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D. M., et al. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behav. Brain Res.* 223, 48–52. doi: 10.1016/j.bbr.2011.03.071

Ward, J. (2004). Emotionally mediated synaesthesia. *Cogn. Neuropsychol.* 21, 761–772. doi: 10.1080/02643290342000393

Zald, D. H., and Pardo, J. V. (2002). The neural correlates of aversive auditory stimulation. *Neuroimage* 16, 746–753. doi: 10.1006/nimg.2002.1115

Zamm, A., Schlaug, G., Eagleman, D., and Loui, P. (2013). Pathways to seeing music: enhanced-structural connectivity in colored-music synesthesia. *Neuroimage* 74, 359–366. doi: 10.1016/j.neuroimage.2013.02.024

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 April 2013; accepted: 05 June 2013; published online: 25 June 2013.

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

Chapter 1, in full, is a reprint of the material as it appears in *Frontiers in Human Neuroscience*. Edelstein, Miren; Brang, David; Rouw, Romke; Ramachandran, V.S. 2013. The dissertation author was the primary investigator and author of this paper.

CHAPTER 2

Context influences how individuals with misophonia respond to certain trigger sounds

Edelstein, M., Monk, B., Rouw, R., and Ramachandran, V.S.

Manuscript in preparation for publication

ABSTRACT

Misophonia is a newly researched condition in which specific sounds cause an intense, aversive response in individuals, characterized by negative emotions and autonomic arousal. Although virtually any sound can become a misophonic "trigger," the most common sounds appear to be bodily sounds related to chewing and eating as well as other repetitive sounds. An intriguing aspect of misophonia is the fact that many misophonic individuals report that they are triggered more, or even only, by sounds produced by specific individuals, and less, or not at all, by sounds produced by animals (although there are always exceptions)

In general, anecdotal evidence suggests that misophonic triggers involve a combination of sound stimuli and contextual cues. The aversive stimulus is more than just a sound and can be thought of as a Gestalt of features which includes sound as a necessary component as well as additional contextual information. In this study, we explore how contextual information influences misophonic responses to human chewing, as well as sonically similar sounds produced by non-human sources. The current study revealed that the exact same sound can be perceived as being much more or less aversive depending on the contextual information presented alongside the auditory information. The results of this study provide a foundation for potential cognitive based therapies.

INTRODUCTION

Misophonia is a newly researched condition in which specific sounds evoke an intensely aversive reaction in sufferers. Misophonia was first described by Jastreboff and Jastreboff (2001) nearly two decades ago but has only recently become a topic of interest to researchers in scientific and clinical communities. Sounds that evoke an intensely aversive reaction in individuals with misophonia are known as “triggers.” When exposed to these trigger sounds, individuals with misophonia experience a variety of physiological and negative emotional responses, resembling a fight-or-flight response (Edelstein et al., 2013; Brout et al., 2018; Kumar et al., 2014). At its most severe, misophonia can be so debilitating that it will often dictate the lives of those who suffer from it, causing people to go to great lengths just to avoid being exposed to certain sounds. Misophonic trigger sounds are frequently sounds that are not regarded as traditionally aversive to most individuals (although they may be considered annoying), and instead are commonly found to be human bodily noises (such as chewing, lip smacking, breathing or sniffing), or other repetitive sounds (such as tapping or pen clicking) (Schröder et al., 2013, Edelstein et al., 2013). While certain trigger sounds (such as chewing and mouthy sounds) appear to be far more common than others, it is important to note that each individual with misophonia possesses their own unique set of trigger sounds and that seemingly any sound has the potential to become a trigger.

When exposed to trigger sounds, misophonic individuals report experiencing intense feelings of anger, anxiety, disgust or rage (Schröder et al., 2013) in addition to a variety of physical sensations such as increased heart rate, tensing of muscles or perceived pressure building up in the body (Edelstein et al., 2013). It has been shown that, in response to auditory stimuli, including trigger sounds, misophonic individuals experience larger physiological responses (SCR and heart rate) indicative of autonomic nervous system arousal, than matched control participants (Edelstein et al., 2013; Kumar et al., 2017).

To date, only two published studies have explored the neural correlates associated with misophonia. Schröder et al. (2014) utilized electroencephalography (EEG) to measure auditory event related potentials (ERPs) in misophonic and control participants during an oddball task. They found that in response to oddball tones, misophonic but not control participants exhibited a decreased mean peak amplitude of the auditory N1 component, which is a component associated with early attention and detecting sudden changes in sensory information. As a decreased N1 component has been observed in individuals with a number of psychiatric conditions, the authors suggest that it could be interpreted as a marker of pathology and that misophonic individuals may be experiencing basic deficits in auditory processing. A groundbreaking study by Kumar et al. (2017) utilized neuroimaging techniques to highlight structural as well as functional neurological differences in those with and without misophonia. Findings revealed that in response to trigger sounds, misophonic participants showed increased activation in the bilateral anterior insular cortex (AIC) as well as increased functional connectivity between the AIC and regions of the brain associated with processing and regulating emotions. As the AIC is thought to be involved in the detection of important, salient stimuli, the increased activation found in misophonic individuals in response to trigger sounds suggests that these sounds are processed as being highly salient. In terms of structural differences, misophonic but not control participants were found to have increased myelination in the ventromedial prefrontal cortex (vmPFC), a region of the brain also involved in regulating emotions.

The prevalence of misophonia in the general population is not well understood yet. In a sample of 483 undergraduate students from a North American university, Wu et al. (2014) found that 20% reported experiencing symptoms of misophonia that were considered clinically significant. Additionally, a study by Zhou et al. (2017) which investigated the prevalence of misophonia in 415 students at two Chinese universities, found that while 16.6% reported clinically significant symptoms, only 6% were classified as experiencing significant levels of impairment. While these studies have made important contributions to our early understanding of misophonia, additional large-scale studies that sample a variety of populations are needed in order to gain an accurate sense of the true prevalence of the condition.

Misophonia has been found to be comorbid with conditions such as obsessive compulsive disorder (OCD) (Schröder et al., 2013; Wu et al., 2014; Ferreira et al., 2013), post-traumatic stress disorder (PTSD) (Rouw & Erfanian, 2017), depression (Wu et al., 2014), generalized anxiety disorder (Ferreira et al., 2013), ADHD (Rouw & Erfanian, 2017), Tourette's syndrome (Neal & Cavanna, 2013), eating disorders (Kluckow et al., 2014) as well as tinnitus and hyperacusis (Jastreboff & Jastreboff, 2014). However, a significant number of individuals with misophonia report that they do not suffer from any additional conditions (Rouw & Erfanian, 2017). More research in this area is needed as there is currently no demonstrable evidence that a relationship exists between misophonia and other conditions (Potgieter et al., 2019).

A number of potential treatments for misophonia have been explored, including cognitive behavioral therapy (CBT) (Schröder et al., 2017; Bernstein et al., 2013; McGuire et al., 2015), tinnitus retraining therapy (TRT) (Jastreboff & Jastreboff, 2014), counterconditioning (Dozier, 2015), mindfulness and acceptance based approaches (Schneider & Arch, 2017) and pharmacological treatment (Vidal et al., 2017; Tunç et al., 2017). However, in addition to varying levels of effectiveness, there looms a significant problem in that these proposed treatments for misophonia are extremely preliminary and have not yet been validated through rigorous scientific testing (Potgieter et al., 2019).

In the last five years, misophonia has often been compared with another emerging sensory phenomenon called the autonomous sensory meridian response (ASMR) in which individuals experience pleasant tingling sensations (usually centralized around the scalp and neck) and feelings of relaxation in response to specific auditory and visual stimuli (Barratt & Davis, 2015; Janik McErlean & Banissy, 2018; Cash et al., 2018). ASMR inducing sounds (also termed "triggers") often include whispering, quiet repetitive noises, crinkling, crisp sounds and sounds indicative of receiving personal attention. Interestingly, many ASMR triggers share striking similarities with misophonic triggers. Additionally, nearly half of the 300 misophonic participants in a study conducted by Rouw & Erfanian (2017) reported experiencing ASMR to certain sounds, suggesting a potential overlap between ASMR and misophonia that should undoubtedly be investigated further.

Despite a growing interest in misophonia in recent years, there still remains a marked lack of empirical research studies investigating the condition. The current study investigates an intriguing characteristic of misophonia reported by Edelstein et al. (2013) that may have the potential to inform future therapies. Namely, many sufferers have reported that sounds produced by certain individuals (typically family members and friends) are particularly aversive, while the same type of sound produced by another individual or a stranger may evoke less of a negative response or none at all. Also, self-produced trigger sounds rarely appear to evoke an aversive response in misophonic individuals. Given that an individual's misophonia often appears to be localized around specific individuals, it seems like the misophonic response could be context sensitive. It has also been reported that the sounds of animals or babies are typically not found to be as aversive as similar sounding trigger sounds produced by adult humans. Although there are always exceptions, based on the aforementioned reports, it appears that an aversive stimulus often involves a highly nuanced formulation of sound and context, suggesting that a misophonic trigger is more than just a sound and instead, a Gestalt of features which includes sound (real or anticipated) as a necessary component. The idea that any singular feature of an aversive stimulus does not necessarily produce aversion on its own, is very interesting and warrants further exploration for both understanding misophonia on a fundamental level, and for its potential for clinically informative results.

Through the use of self-reported aversiveness ratings, we assessed participant aversion to a variety of classic trigger sounds in the presence and absence of contextual information. Clips of common trigger sounds (crunchy/wet human eating sounds) as well as sounds that highly resembled trigger sounds (crunchy/wet animal eating sounds and various crunchy/wet non eating sounds) were presented to self-identified misophonic and age/gender matched control participants in three experimental blocks. In each of the three experimental blocks, the type of contextual information accompanying each sound differed slightly. In block 1, participants were presented with only the audio of the sounds, and not given any feedback about what they were listening to. In block 2, participants were also presented with only the audio of the sounds, but prior to each sound, received a short text description about what they were potentially listening to. However, participants were informed that this description was not always correct and it was up to them to decide if the description matched the sound presented. In block 3, participants were presented with both the audio and video of each sound, which ultimately revealed the identity of each sound they had been listening to.

By utilizing deliberately ambiguous sounds and manipulating the type of contextual information provided about said sounds, our intention was to influence what participants believed they were listening to, to the extent where they may be convinced that certain trigger sounds were actually non-trigger sounds and certain non-trigger sounds were actually trigger sounds, and observe if their beliefs influenced their reactions. We hypothesized that misophonic individuals (but not controls) would find sounds that they perceived to be human eating sounds (regardless of whether they actually were or not) to be significantly more aversive than sounds that they perceived to be animal eating and non eating sounds. If successful, this study would demonstrate that contextual information that an individual associates with a sound can significantly influence their response to that sound, providing empirical evidence for the idea that the physical properties of a trigger sound are not the only factors driving the misophonic response.

METHODS

Participants:

Twenty self-identified misophonic participants (5 males and 15 females; mean age = 30.4 years; range = 20-58) and twenty age and gender matched control participants (5 males and 15 females; mean age = 31.24 years; range = 20-58) were recruited from the student population at the University of California, San Diego and the greater San Diego area. All participants reported normal vision and hearing and signed a consent form approved by the UCSD Human Research Protections Program prior to participating. Participants were reimbursed with either UCSD course credit or at a rate of $10/hour. The entire lab session lasted for approximately 2 hours.

Questionnaires:

Control participants filled out a short demographic form that also assessed any prior knowledge of misophonia and sought to determine whether they may suffer from the condition unknowingly. No control participants were found to experience misophonic symptoms. Self-identified misophonic participants were given a demographic form as well as several commonly used misophonia questionnaires that assessed their experiences with the condition and gauged the severity of their symptoms. The questionnaires included were the Amsterdam Misophonia Scale (A-MISO-S), which measures the severity of the symptoms and intensity of responses associated with a participant's misophonia, the Misophonia Activation Scale (MAS-1) which characterizes eleven levels (0-10) of misophonia severity, and the Misophonia Assessment Questionnaire (MAQ) which assesses how frequently participants experience negative effects and disturbances associated with misophonia.

Misophonic participants scored an average of 11.7 (range = 7-24) points out of a maximum of 24 (most severe) points on the A-MISO-S and an average of 28 (range: 10-63) points out of a maximum of 63 points (most severe) on the MAQ. Of the eleven levels of misophonia severity detailed in the MAS-1 (0-10), the average level amongst participants was found to be 5.475 (range = 3.5-9).

Experimental Setup:

As a general overview, each participant took part in a session that consisted of 3 experimental blocks. Although it differed slightly from block-to-block, the general structure of a block was as follows: participants were seated 20 inches away from a computer screen and wore a pair of Sennheiser headphones. Through the use of MATLAB R2014B, visual stimuli were presented on the computer screen and auditory stimuli were presented through the headphones at 50% of the computer's volume. An individual trial consisted of a 5 second (pre-stimulus period) followed by a 15 second clip (stimulus period), and finally a 10 second intertrial interval (ITI). During the ITI, participants were instructed to verbally make an aversiveness rating about the clip they were just presented with on a 1-10 scale. Each block contained 36 clips, with each clip falling into one of three sound categories (Fig. 1).

Participants were informed beforehand that an aversiveness rating of "1" signified very little to no discomfort while a rating of "10" signified extreme discomfort and possibly a strong desire to leave the room should the sound continue. Each aversiveness rating was recorded by the experimenter. Between blocks, participants were instructed to take a short break.

Stimuli:

Thirty-six, 15-second video clips were used in the study. All clips were either found on Youtube or created in the lab. Each clip was placed into one of three sound categories: human eating, animal eating or non eating, with 12 clips in each category. Clips were selected based on the criteria that they either were or highly resembled classic misophonic trigger sounds (most were crunchy or wet sounding in nature). Audio (sound only) and audio-visual (sound + video) versions of each clip were created.

Clips were selected based on results from a pilot study involving 21 participants that was conducted in the summer of 2016. The purpose of this pilot study was to identify a set of classic misophonic sounds that could plausibly be interpreted as belonging to more than one of the aforementioned sound categories (when presented with only audio and no visuals). The most categorically ambiguous clips were then selected to be used as stimuli in the current study.

Experimental Blocks:

For each block 1 trial, participants were presented with a 5 second pre-stimulus period followed by a 15 second audio only clip, and then a 10 second ITI during which they made their aversiveness rating on a 1-10 scale. In addition to their aversiveness rating, participants were also instructed to make a guess as to what they thought the sound source of each clip was (based on the aforementioned 3 sound categories) during this ITI. The sounds in block 1 were presented in a randomized order for every participant (Fig. 1).

For each block 2 trial, participants were presented with a 5 second pre-stimulus period which included 1 second of blank, black screen, followed by 3 seconds of descriptive text, followed by 1 second of blank, black screen. Next came a 15 second audio only clip and then the 10 second ITI during which participants made their aversiveness rating on a 1-10 scale. Half of the time the text presented during the pre-stimulus period was a correct description of the sound that would play immediately after it and half of the time it was an incorrect description (Fig. 1). When it was incorrect, the text was a randomly selected description from one of the other two sound categories that the sound from that trial did not fall under. An incorrect description was never from the same category as the sound presented. For each trial in block 2, participants responded with a "yes" or "no" as to whether or not the text description they received sounded like the sound they were presented with. Participants were instructed to make this judgment based on general sound category and not the specifics of each description. The ordering of both the textual descriptions and sounds were preselected for each participant and counterbalanced. There were 6 possible sound and 6 possible text pseudo-random orderings. Each misophonic participant was matched with a control participant who received the same sound and text ordering in block 2 as they did.

For each block 3 trial, participants were presented with a 5 second pre-stimulus period followed by a 15 second video clip (audio and video) and then a 10 second ITI during which participants made their aversiveness rating on a 1-10 scale. Video clips in block 3 were presented in a randomized order for each participant (Fig. 1).

RESULTS

Within Blocks Results

<u>Block 1: Audio Only | All Trials:</u>

As expected, we found that overall, aversiveness ratings given by misophonics ($M = 4.92$) were significantly higher than ratings given by controls ($M = 1.97$) [$F(1,38) = 49.764$, $p < .001$]. There was also an observed within subject factor effect of Sound Category [$F(2,76) = 25.719$, $p < .001$], where aversion to sounds from the human eating category ($M = 4.01$) was significantly higher than aversion to sounds from the animal eating ($M = 3.18$) and non eating ($M = 3.15$) categories, across groups (Fig. 2A).

This observed main effect of Sound Category was driven by a significant interaction between Group and Sound Category [$F(2,76) = 21.406$, $p < .001$], where misophonic participants rated human eating sounds as particularly aversive compared with animal eating and non eating sounds ($M = 5.98, 4.55, 4.25$ respectively for misophonics; $M = 2.05, 1.82, 2.06$ respectively for controls). As a follow up to the interaction, paired t-tests indicated that misophonic participants rated human eating sounds as significantly more aversive than both animal eating [$t(19) = 6.36$, $p < .001$] and non eating [$t(19) = 5.27$, $p < .001$] sounds (there was no statistical difference between animal eating vs non eating sounds [$t(19) = 1.563$, $p = .135$]) (Fig. 2A).

Figure 4 depicts average misophonic and control ratings of each stimulus from block 1 in the form of a scatterplot, illustrating the finding that misophonic participants found all stimuli to be more aversive than controls did as well as showing which sounds were found to be most aversive.

Block 1: Audio Only | Correct Trials:

For ratings of trials where participants correctly identified the sound category, we found significant main effects of Group [$F(1,38) = 48.109$, $p < .001$], with the misophonic group rating sounds as significantly more aversive overall ($M = 4.91$) than the control group ($M = 1.97$), and Sound Category [$F(2,76) = 28.552$, $p < .001$], with human eating sounds rated as more aversive ($M = 4.22$) than animal eating ($M = 3.26$) and non eating sounds ($M = 2.83$) across groups. A significant interaction between Group and Sound Category [$F(2,76) = 19.801$, $p < .001$] was also observed, with human eating sounds rated as particularly aversive by the misophonic participants ($M = 6.28$), compared with animal eating ($M = 4.75$) and non eating sounds ($M = 3.71$) (Fig. 2C).

Paired t-tests confirmed that misophonics rated human eating sounds as significantly more aversive than animal eating sounds [$t(19) = 5.09$, $p < .001$] and non eating sounds [$t(19) = 5.60$, $p < .001$]. Interestingly, misophonics also rated animal eating sounds as significantly more aversive than non eating sounds [$t(19) = 3.79$, $p = .001$]. Controls were found to rate human eating sounds as significantly more aversive than animal eating sounds, [$t(19) = 3.13$, $p = .006$], but not non eating sounds [$t(19) = 1.50$, $p = .149$]. Controls also found non eating sounds to be marginally more aversive than animal eating sounds [$t(19) = 1.826$, $p = .084$].

Block 1: Audio Only | Incorrect Trials:

For ratings of trials where the participant incorrectly identified the sound category, we found a significant main effect of Group [$F(1,38) = 42.685$, $p < .001$], with the misophonic group rating sounds as significantly more aversive overall (M = 4.59) than the control group (M = 1.94), a marginally significant main effect of Sound Category [F(2,76) = 2.557, p = .084], (human eating sounds (M = 3.21), animal eating (M = 3.08), non eating sounds (M = 3.51) across groups), but no significant interaction between Group and Sound Category [F(2,76) = .593, p = .475], with human eating sounds not rated as particularly aversive by the misophonic participants (M = 4.65) when compared with animal eating (M = 4.28) and non eating sounds (M = 4.84) (Fig. 2B).

Paired t-tests indicated that misophonics did not demonstrate any significant difference in aversiveness ratings between incorrectly identified human eating sounds and incorrectly identified animal eating sounds [t(19) = .938, p = .36] or incorrectly identified non eating sounds [t(19) = -.59, p = .562]. There was also no significant difference between the ratings of incorrectly identified animal eating sounds and incorrectly identified non eating sounds [t(19) = -1.67, p = .112] for misophonics. Controls did not demonstrate any significant difference in aversiveness ratings between incorrectly identified human eating sounds and incorrectly identified animal eating sounds [t(19) = -.892, p = .384] or incorrectly identified non eating sounds [t(19) = -1.80, p = .088]. There was also no significant difference between the ratings of incorrectly identified animal eating sounds and incorrectly identified non eating sounds [t(19) = -1.582, p = .130] for controls.

<u>Block 1: Audio Only | Trial Comparisons:</u>

Additionally, the ratings of trials where the participant incorrectly identified the sound category were compared to trials where they correctly identified the sound category. Specifically, we were interested in comparing 1) ratings of trials where human eating sounds were misidentified as either animal eating sounds or non eating sounds to ratings trials where human eating sounds were correctly identified as human eating sounds, 2) ratings of trials where animal eating sounds were misidentified as human eating sounds to ratings of trials where animal eating sounds were correctly identified as animal eating sounds and 3) ratings of trials where non eating sounds were misidentified as human eating sounds to ratings of trials where non eating sounds were correctly identified as non eating sounds.

Results showed that misophonics rated human eating sounds incorrectly identified as animal eating or non eating sounds ($M = 4.64$, $SD = 2.26$) as significantly less aversive than human eating sounds correctly identified as human eating sounds ($M = 6.28$, $SD = 2.18$), [$t(19) = -4.46$, $p < .001$]. The same pattern was present for controls [$t(19) = -2.2$, $p = .04$]. Although there was no significant difference between misophonic ratings of animal eating sounds incorrectly identified as human eating sounds ($M = 5.04$, $SD = 1.7$) and animal eating sounds correctly identified as animal eating sounds ($M = 4.75$, $SD = 1.97$), [$t(19) = -.865$, $p = .398$]s, controls did show a significant difference in ratings between these two groups of trials, [$t(19) = -2.25$, $p = .036$]. Lastly, misophonics rated non eating sounds incorrectly identified as human eating sounds ($M = 5.2$, $SD = 1.41$) as significantly more aversive than non eating sounds correctly identified as non eating sounds ($M = 3.7$, $SD = 1.91$), [$t(19) = 3.08$, $p = .006$]. Controls also exhibited the same pattern of results for non eating sounds [$t(19) = 3.3$, $p = .004$] (Fig. 2C).

Block 1: Audio Only: Stimulus Classification | Category Guess Propensity & Accuracy:

We also investigated the level of accuracy for sound category identification (percentage of trials correct) with factors of Group (misophonics, controls) and Sound Category (human eating, animal eating, non eating). No significant main effect of Group [$F(1,38) = .615$, $p = .438$]

was observed, but there was a significant main effect of Sound Category [F(2,76) = 18.156, p < .001] as well as a marginally significant interaction between Group and Sound Category [F(2,76) = 3.353, p = .04]. This interaction brings about a few interesting findings. The first finding revealed that misophonic participants (M = 79.175%, SD = 16.120%) were significantly more accurate than controls (M = 67.075%, SD = 12.8293%) when identifying human eating sounds in particular [F(1,38) = 6.899 , p = .012] but not when identifying animal eating sounds [F(1,38) = .034, p = .854] or non eating sounds [F(1,38) = 1.756, p = .193] (Fig. 3A).

Paired t-tests indicated that misophonics were also significantly more accurate at identifying human eating sounds than animal eating sounds [t(19) = 5.361, p < .001], significantly more accurate at identifying non eating sounds than animal eating sounds [t(19) = 2.482, p = .023] and marginally more accurate at identifying human eating sounds than non eating sounds [t(19) = 1.885, p = .075]. Additionally, controls were significantly more accurate at identifying human eating sounds than animal eating sounds, [t(19) = 4.708, p < .001], significantly more accurate at identifying non eating sounds than animal eating sounds [t(19) = 4.162, p = .001] but not significantly more accurate at identifying non eating sounds than human eating sounds [t(19) = 1.256, p = .224].

In order to address the possibility that participants may have demonstrated a preference to make guesses within a specific sound category (which could influence their accuracy), the percentage of trials that were guessed to be in each sound category was investigated with factors of Group (misophonics, controls) and Sound Category (human eating, animal eating, non eating). Although no significant main effects of Group [F(1,38) = .580, p = .451] or Sound Category [F(2,76) = 1.9, p = .157] were observed, a significant interaction between Group and Sound Category [F(2,76) = 5.472, p = .006] was found. This interaction brings about a few interesting findings. The first finding revealed that misophonic participants were significantly more likely than controls to guess that a sound was a human eating sound [$F(1,38) = 9.37$, $p = .004$] but not significantly more likely than controls to guess that a sound was an animal eating sound [$F(1,38) = .233$, $p = .632$]. Interestingly, controls were significantly more likely than misophonics to guess that a sound was a non eating sound [$F(1,38) = 5.395$, $p = .026$] (Fig. 3B).

Paired t-tests indicated that misophonics were also significantly more likely to guess that a sound was a human eating sound as opposed to an animal eating sound [$t(19) = 2.667$, $p = .015$], or a non eating sound [$t(19) = 2.16$, $p = .044$]. No significant difference in guessing rate was found between animal eating and non eating sounds [$t(19) = .164$, $p =.871$]. Additionally, although controls were not significantly more likely to guess that a sound was a human eating sound as opposed to an animal eating sound [$t(19) = .202$, $p = .842$], they were marginally more likely to guess that a sound was a non eating sound as opposed to a human eating sound [$t(19) = 1.936$, $p = .068$], or an animal eating sound [$t(19) = 2.011$, $p = .059$].

<u>Block 2: Audio + Text | Agree + Disagree Trials:</u>

First, we conducted a repeated measures mixed design ANOVA on factors of Group (misophonic, control) and Sound Category (human eating, animal eating, non eating) for ratings of all block 2 trials (regardless of whether the participant received an accurate (target) or false (foil) textual description and regardless of whether they got the trial right or wrong). Overall, we found significant main effects of Group [$F(1,38) = 51.1$, $p < .001$] and Sound Category [$F(2,76) = 34.7$, $p < .001$] as well as a significant interaction between Group and Sound Category [$F(2,76) = 27.6$, $p < .001$].

Follow up paired t-tests revealed that misophonics rated human eating sounds (M = 6.14, SD = 2.03) as significantly more aversive than both animal eating (M = 5.00, SD = 2.01), [t(19) = 5.87, p < .001] and non eating sounds (M = 4.45, SD = 1.8), [t(19) = 6.69, p <.001]. Additionally, misophonics rated animal eating sounds as significantly more aversive than non eating sounds [t(19) = 3.37, p = .003]. Controls demonstrated a similar pattern of results and rated human eating sounds (M = 2.1, SD = .79) as significantly more aversive than animal eating sounds (M = 1.87, SD = .6), [t(19) = 3.301, p = .004] but not non eating sounds (M = 2.05, SD = .719) [t(19) = .669, p = .512]. Additionally, controls rated non eating sounds as significantly more aversive than animal eating sounds [t(19) = 2.75, p = .013] (Fig. 5A).

Block 2: Audio + Text | Agree Trials:

We compared the ratings of trials where participants incorrectly believed false text (foil) descriptions preceding the stimulus to trials where they correctly believed true text (target) descriptions preceding the stimulus. Specifically, we were interested in comparing 1) ratings of trials where human eating sounds were incorrectly believed to be either animal eating sounds or non eating sounds to ratings of trials where human eating sounds were correctly believed to be human eating sounds, 2) ratings of trials where animal eating sounds were incorrectly believed to be human eating sounds to ratings of trials where animal eating sounds were correctly believed to be animal eating sounds and 3) ratings of trials where non eating sounds were incorrectly believed to be human eating sounds to ratings of trials where non eating sounds were correctly believed to be non eating sounds.

Specifically, misophonics rated human eating sounds incorrectly believed to be animal eating or non eating sounds (M = 4.83, SD = 1.81) as significantly less aversive than human eating sounds correctly believed to be human eating sounds (M = 6.59, SD = 2.23), [t(19) = -4.344, p < .001]. Misophonics also rated animal eating sounds incorrectly believed to be human eating sounds (M = 5.56, SD = 2.00) as significantly more aversive than animal eating sounds correctly believed to be animal eating sounds (M = 4.79, SD = 2.21), [t(19) = 1.69, p = .05]. Additionally, misophonics rated non eating sounds incorrectly believed to be human eating sounds (M = 5.58, SD = 1.55) as significantly more aversive than non eating sounds correctly believed to be eating sounds (M = 4.23, SD = 1.92), [t(19) = 3.03, p = .0035] (Fig. 5B).

Controls did not rate animal eating sounds incorrectly believed to be human eating sounds as significantly more aversive than animal eating sounds correctly believed to be animal eating sounds [p >.05]. They also did not rate non eating sounds incorrectly believed to be human eating sounds as significantly more aversive than non eating sounds correctly believed to be eating sounds [p >.05]. However, controls did rate human eating sounds incorrectly believed to be animal eating or non eating sounds (M = 1.78, SD = .76) as significantly less aversive than human eating sounds correctly believed to be human eating sounds (M = 2.31, SD = 1.07), [t(19) = -2.13, p = .047] (Fig. 5B).

<u>Block 3: Audio + Video Trials:</u>

We conducted a repeated measures mixed design ANOVA on factors of Group (misophonic, control) and Sound Category (human eating, animal eating, non eating) for ratings of block 3 trials. Overall, we found significant main effects of Group [F(1,38) = 46.822, p < .001] and Sound Category [F(2,76) = 51.879, p < .001] as well as a significant interaction between Group and Sound Category [F(2,76) = 21.081, p < .001].

Follow up paired t-tests revealed that misophonics rated human eating sounds (M = 6.97, SD = 2.01) as significantly more aversive than both animal eating (M = 3.99, SD = 2.28), [t(19) = 7.39, p < .001] and non eating sounds (M = 6.57, SD = 2.08), [t(19) = 6.69, p < .001]. Misophonics did not rate animal eating sounds as significantly more aversive than non eating sounds [t(19) = -.186, p = .854]. Controls rated human eating sounds (M = 2.44, SD = .87) as significantly more aversive than animal eating sounds (M = 1.63, SD = .46), [t(19) = 5.134, p < .001] and non eating sounds (M = 1.94, SD = .63) [t(19) = 3.13, p = .006]. Additionally, controls rated non eating sounds as significantly more aversive than animal eating sounds [t(19) = 3.29, p = .004] (Fig. 6).

Between Blocks Results

In addition to investigating how misophonic and control participants responded to human eating, animal eating and non eating sounds within the differing contexts of blocks 1, 2 and 3, we also examined how their responses to specific sounds changed across these blocks. In particular, we were interested in observing how their ratings changed between all three blocks for

1) human eating sounds that were correctly identified as human eating sounds in block 1 but were believed to be produced by nonhuman (animal eating and non eating sounds) sources in block 2.

2) nonhuman sounds (animal eating and non eating sounds) that were correctly identified as nonhuman sounds in block 1 but were believed to be human eating sounds in block 2.

3) human eating sounds that were correctly identified as human eating sounds in blocks 1 and 2.

4) nonhuman sounds (animal eating and non eating sounds) that were correctly identified as nonhuman sounds in blocks 1 and 2.

When considering human eating sounds that were correctly identified as human eating sounds in block 1 but were believed to be produced by nonhuman sources in block 2, we find that misophonics rated these sounds to be significantly more aversive in block 1 ($M = 5.95$, $SD = 1.69$) than when they encountered them again in block 2 ($M = 5.0$, $SD = 1.58$), [$t(13) = 1.892$, $p = .04$]. Misophonics additionally rated these sounds to be significantly more aversive in block 3 ($M = 6.45$, $SD = 1.45$) than block 2 ($M = 5.0$, $SD = 1.58$), [$t(13) = 3.065$, $p = .0045$], but no significant difference in ratings for these sounds was found between blocks 1 and 3, [$p > .05$]. Controls did not exhibit significant differences in ratings for these sounds between blocks 1 and 2 [$p > .05$], blocks 2 and 3 [$p > .05$] or blocks 1 and 3 [$p > .05$] (Fig. 7A).

When considering nonhuman sounds (animal eating and non eating sounds) that were correctly identified as nonhuman sounds in block 1 but were believed to be human eating sounds in block 2, we find that misophonics rated these sounds to be significantly more aversive in block 2 ($M = 5.7$, $SD = 2.32$) than in block 1 ($M = 4.14$, $SD = 2.5$), [$t(19) = 4.098$, $p < .001$]. Misophonics also rated these sounds as significantly more aversive in block 2 ($M = 5.7$, $SD = 2.32$) than in block 3 ($M = 3.53$, $SD = 2.26$), [$t(19) = 4.875$, $p < .001$], but no significant difference in ratings for these sounds was found between blocks 1 and 3, [$p > .05$]. Controls did not exhibit significant differences in ratings for these sounds between blocks 1 and 2 [$p > .05$] or blocks 1 and 3 [$p > .05$], but a significant difference between blocks 2 and 3 was observed, with these sounds being rated as significantly more aversive in block 2 ($M = 2.08$, $SD = 1.03$) than block 3 ($M = 1.59$, $SD = .712$), [$t(16) = 3.125$, $p = .007$] (Fig. 7B).

When comparing human eating sounds that misophonics correctly identified as human eating sounds in block 1 but were believed to be produced by nonhuman sources in block 2, with nonhuman sounds (animal eating and non eating sounds) that misophonics correctly identified as nonhuman sounds in block 1 but were believed to be human eating sounds in block 2, we find a significant main effect of the between subject factor of Sound Type [$F(1,32) = 4.57$, $p = .04$] and a significant interaction between Sound Type and the within subject factor Block [$F(2,64) = 17.254$, $p < .001$] (Fig. 7C).

When considering human eating sounds that were correctly identified as human eating sounds in blocks 1 and 2, we find that misophonics rated these sounds to be marginally more aversive in block 2 ($M = 6.6$, $SD = 2.29$) than when they encountered them in block 1 ($M = 6.32$, $SD = 2.36$), [$t(19) = 1.48$, $p = .08$], significantly more aversive in block 3 ($M = 7.07$, $SD = 2.15$) than block 2 ($M = 6.6$, $SD = 2.29$), [$t(19) = 2.33$, $p = .015$], and significantly more aversive in block 3 ($M = 7.07$, $SD = 2.15$) than block 1 ($M = 6.32$, $SD = 2.36$), [$t(19) = 3.28$, $p = .002$]. Controls did not exhibit significant differences in ratings for these sounds between blocks 1 and 2 [$p > .05$], blocks 2 and 3 [$p > .05$] or blocks 1 and 3 [$p > .05$] (Fig. 8A).

When considering nonhuman sounds (animal eating and non eating sounds) that were correctly identified as nonhuman sounds in blocks 1 and 2, we find that misophonics did not exhibit significant differences in ratings for these sounds between blocks 1 and 2 [$p > .05$], blocks 2 and 3 [$p > .05$] or blocks 1 and 3 [$p > .05$]. Controls also did not exhibit significant differences in ratings for these sounds between blocks 1 and 2 [$p > .05$], blocks 2 and 3 [$p > .05$] or blocks 1 and 3 [$p > .05$] (Fig. 8B).

When comparing human eating sounds that misophonics correctly identified as human eating sounds in blocks 1 and 2 and nonhuman sounds (animal eating and non eating sounds) that misophonics correctly identified as nonhuman sounds in blocks 1 and 2, we find a significant main effect of the between subject factor of Sound Type [$F(1,56) = 14.53$, $p < .001$] and a significant interaction between Sound Type and the within subject factor Block [$F(2,112) = 3.42$, $p = .036$] (Fig. 8C).

Lastly, when investigating differences in how individual stimuli were rated in block 3 and block 1, we find that misophonic participants had a larger range of difference scores overall than controls. Additionally, although both misophonic and control participants tended to rate human eating sounds as more aversive in block 3 than block 1, and animal eating sounds as less aversive in block 3 than block 1, misophonics demonstrated this to a much greater extent (Fig. 9).

DISCUSSION

Overall these results support our main hypothesis that context plays a role in how aversive misophonic participants find certain sounds to be. This is in line with previous reports that suggest that a sound's source is a crucial factor in determining what is considered a misophonic trigger sound to an individual (Edelstein et al., 2013; Schneider & Arch, 2017). However, the findings from our experiment extend beyond this and show that while sound source is indeed an important factor in the misophonic response, an individual's *perception* of a sound's source is enough to influence how they respond to that sound.

Our hypothesis that misophonic participants would find human eating sounds as most aversive when compared with animal eating and non eating sounds was confirmed by within block analyses of ratings and skin conductance of blocks 1, 2 and 3. In block 1, we showed that in the absence of any contextual information (such as text description or video), whether or not a participant correctly guessed a sound's source (and specifically what they thought the sound was when they didn't guess correctly), played a role in how aversive they rated that sound to be. Block 2 showed a similar finding, where correct or incorrect text descriptions provided prior to each sound (and whether or not participants believed these descriptions), influenced how aversive participants found those sounds to be. Block 3, which included video of the sound participants were listening to, left no room for interpretation and further solidified the finding that human eating sounds were considered by both misophonic and control participants to be significantly more aversive than animal eating and non eating sounds. Although both groups found human eating sounds to be the most aversive sound category, misophonic individuals always showed much higher aversiveness ratings than controls overall.

In addition to examining how participants responded to these three categories of sounds within blocks, we examined how responses to specific sounds within these categories may change across blocks. In particular, we found that the very same sound could be rated significantly differently from block to block when paired with different contextual information. Specifically, we were interested in the rating change across blocks of human eating and nonhuman sounds (animal eating and non eating sounds were grouped together to form this category) that were identified correctly in block 1, but were believed to be nonhuman sounds and human eating sounds, respectively, when they were heard again in block 2. Indeed, we found that misophonics, but not controls, rated the very same human eating sounds that were correctly identified in block 1, as significantly less aversive when encountered again in block 2 when believed to be nonhuman sounds. When misophonics encountered those same human eating sounds for a third time in block 3, with video, their ratings significantly increased from block 2. Conversely, we found that misophonics, but not controls, rated the very same nonhuman sounds that were correctly identified in block 1, as significantly more aversive when encountered again in block 2 when believed to be human eating sounds. When misophonics encountered those same nonhuman sounds for a third time in block 3, with video, their ratings significantly decreased from block 2.

We were also interested in the rating change across blocks of human eating and nonhuman sounds that were correctly identified as human eating and nonhuman sounds, respectively, in both blocks 1 and 2. While controls did not exhibit significant differences in ratings between blocks for either of these groups of sounds, misophonics did, but only for human eating sounds and not nonhuman sounds. Specifically, misophonics rated human eating sounds as increasingly aversive from blocks 1 to 3. This suggests that for trigger sounds, such as human eating sounds, the more contextual information misophonics are given about what they were listening to, the more aversive the sound becomes.

In terms of future directions, it would be worthwhile to develop a reliable technique to assess physiological markers of misophonia. It should be noted that we collected SCR and electromyography (EMG) data from the participants in this specific study in order to supplement their subjective aversiveness ratings, but unfortunately, due to a number of factors such as a low signal to noise ratio, outdated equipment and the length of the study, not enough of the physiological data ended up being clean enough to be properly analyzed. However, with higher quality recordings, some of the observed main effects from this study would likely produce reliable physiological components.

Ultimately, the findings from this study demonstrate that sound source plays a large role in what are considered to be trigger sounds. The idea that two sounds could sonically sound very similar to each other, but only one might trigger an individual with misophonia, suggested that there is much more that goes into a misophonic trigger than just the sound itself. Through the exclusive use of sonically similar sounds, this study not only showed that human eating sounds were considered to be significantly more aversive than animal eating and non eating sounds to misophonic individuals overall; it also showed that how one interprets these sounds can significantly influence how aversive they believe them to be. The findings from this study show that, depending on the contextual information given, the very same sound could be considered significantly more or less aversive the next time it was encountered. There is already preliminary evidence that cognitive behavioral therapy, which utilizes techniques to help patients reappraise negative thoughts and feelings, may be helpful for individuals with misophonia (Schröder et al., 2017). The fact that there appears to be some degree of cognitive flexibility in terms of reassessing misophonic trigger sounds leads us to believe that there may be successful therapeutic applications of this work in the future.

References

Barratt, E. L., & Davis, N. J. (2015). Autonomous Sensory Meridian Response (ASMR): a flow-like mental state. *PeerJ*, 3, e851.

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, e10.

Brout, J. J., Edelstein, M., Erfanian, M., Mannino, M., Miller, L. J., Rouw, R., Kumar, S. & Rosenthal, M. Z. (2018). Investigating Misophonia: A Review of the Empirical Literature, Clinical Implications, and a Research Agenda. *Front. Neurosci*. 12:36. doi: 10.3389/fnins.2018.00036

Bruxner, G. (2016). 'Mastication rage': a review of misophonia–an under-recognised symptom of psychiatric relevance?. *Australasian Psychiatry*, 24(2), 195-197.

Cash, D. K., Heisick, L. L., & Papesh, M. H. (2018). Expectancy effects in the Autonomous Sensory Meridian Response. *PeerJ*, 6, e5229.

Dozier, T. H. (2015). Counterconditioning treatment for Misophonia. *Clinical Case Studies*, 14(5), 374–387.

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

Ferreira, G. M., Harrison, B. J., & Fontenelle, L. F. (2013). Hatred of sounds: misophonic disorder or just an underreported psychiatric symptom?. *Annals of clinical psychiatry: official journal of the American Academy of Clinical Psychiatrists*, 25(4), 271-274.

Janik McErlean, A. B., & Banissy, M. J. (2018). Increased misophonia in self-reported Autonomous Sensory Meridian Response. *PeerJ*, 6, e5351.

Jastreboff, M. M., & Jastreboff, P. J. (2001). Components of decreased sound tolerance: hyperacusis, misophonia, phonophobia. *ITHS News Lett.* 2, 5–7.

Jastreboff, P. J., & Jastreboff, M. M. (2014). Treatments for decreased sound tolerance (hyperacusis and misophonia). *Seminars in Hearing* (Vol. 35, No. 02, pp. 105-120).

Kluckow, H., Telfer, J., & Abraham, S. (2014). Should we screen for misophonia in patients with eating disorders? A report of three cases. *International Journal of Eating Disorders*, 47(5), 558-561.

Kumar, S., Hancock, O., Cope, T., Sedley, W., Winston, J., & Griffiths, T. D. (2014). Misophonia: A disorder of emotion processing of sounds. *Journal of Neurology, Neurosurgery & Psychiatry*, 85(8), e3–e3.

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

McGuire, J. F., Wu, M. S., & Storch, E. A. (2015). Cognitive-behavioral therapy for 2 youths with misophonia. *The Journal of Clinical Psychiatry*, 76(5), 573–574.

Neal, M., & Cavanna, A. E. (2013). Selective sound sensitivity syndrome (misophonia) in a patient with Tourette syndrome. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 25(1), E01.

Potgieter, I., MacDonald, C., Partridge, L., Cima, R., Sheldrake, J., & Hoare, D. J. (2019). Misophonia: A scoping review of research. *J. Clin. Psychol*. doi: 10.1002/jclp.22771.

Rouw, R., & Erfanian, M. (2017). A large-scale study of misophonia. *J. Clin. Psychol*. doi: 10.1002/jclp.22500.

Schneider, R. L., & Arch, J. J. (2017). Case study: a novel application of mindfulness-and acceptance-based components to treat misophonia. *Journal of Contextual Behavioral Science*, 6(2), 221-225.

Schröder, A. E., van Diepen, R., Mazaheri, A., Petropoulos-Petalas, D., de Amesti, V., Vulink, N. C., & Denys, D. A. (2014). Diminished n1 auditory evoked potentials to oddball stimuli in misophonia patients. *Front. Behav. Neurosc*i. 8:123. doi: 10.3389/fnbeh.2014.00123

Schröder, A. E., Vulink, N. C., & Denys, D. A. (2013). Misophonia: diagnostic criteria for a new psychiatric disorder. *PLoS ONE* 8:e54706. doi: 10.1371/journal.pone.0054706

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

Tunç, S., & Başbuğ, H. S. (2017). An extreme physical reaction in misophonia: Stop smacking your mouth! *Psychiatry and Clinical Psychopharmacology*, 27(4), 416–418.

Vidal, C., Vidal, L. M., & Lage, M. A. (2017). Misophonia: Case report. *European Psychiatry*, 41, S644.

Wu, M. S., Lewin, A. B., Murphy, T. K., & Storch, E. A. (2014). Misophonia: incidence, phenomenology, and clinical correlates in an undergraduate student sample. *J. Clin. Psychol*. 70, 994–1007. doi: 10.1002/jclp.22098

Zhou, X., Wu, M. S., & Storch, E. A. (2017). Misophonia symptoms among Chinese university students: Incidence, associated impairment, and clinical correlates. *Journal of Obsessive-Compulsive and Related Disorders*, 14,7–12.

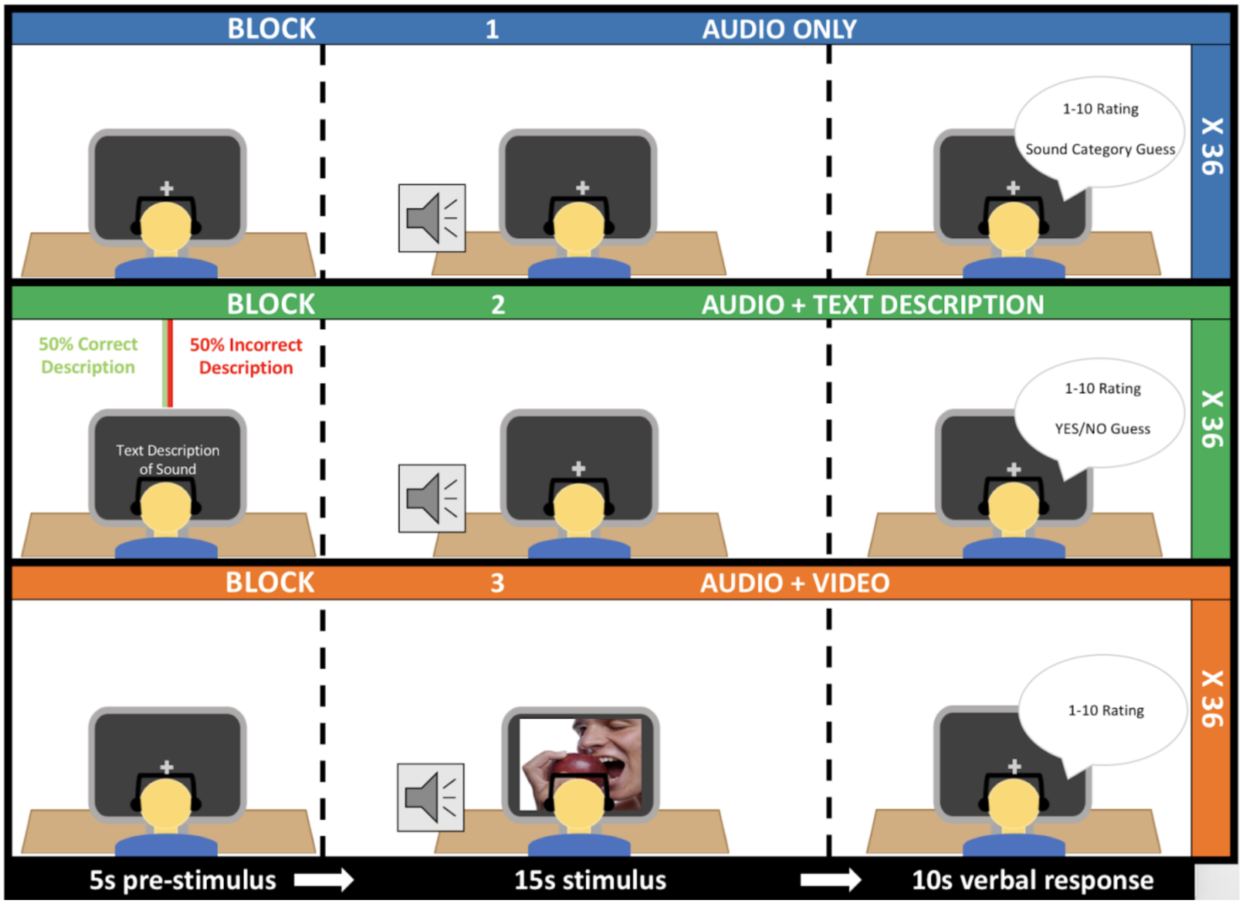

Figure 2.1. Experimental setup and procedure.

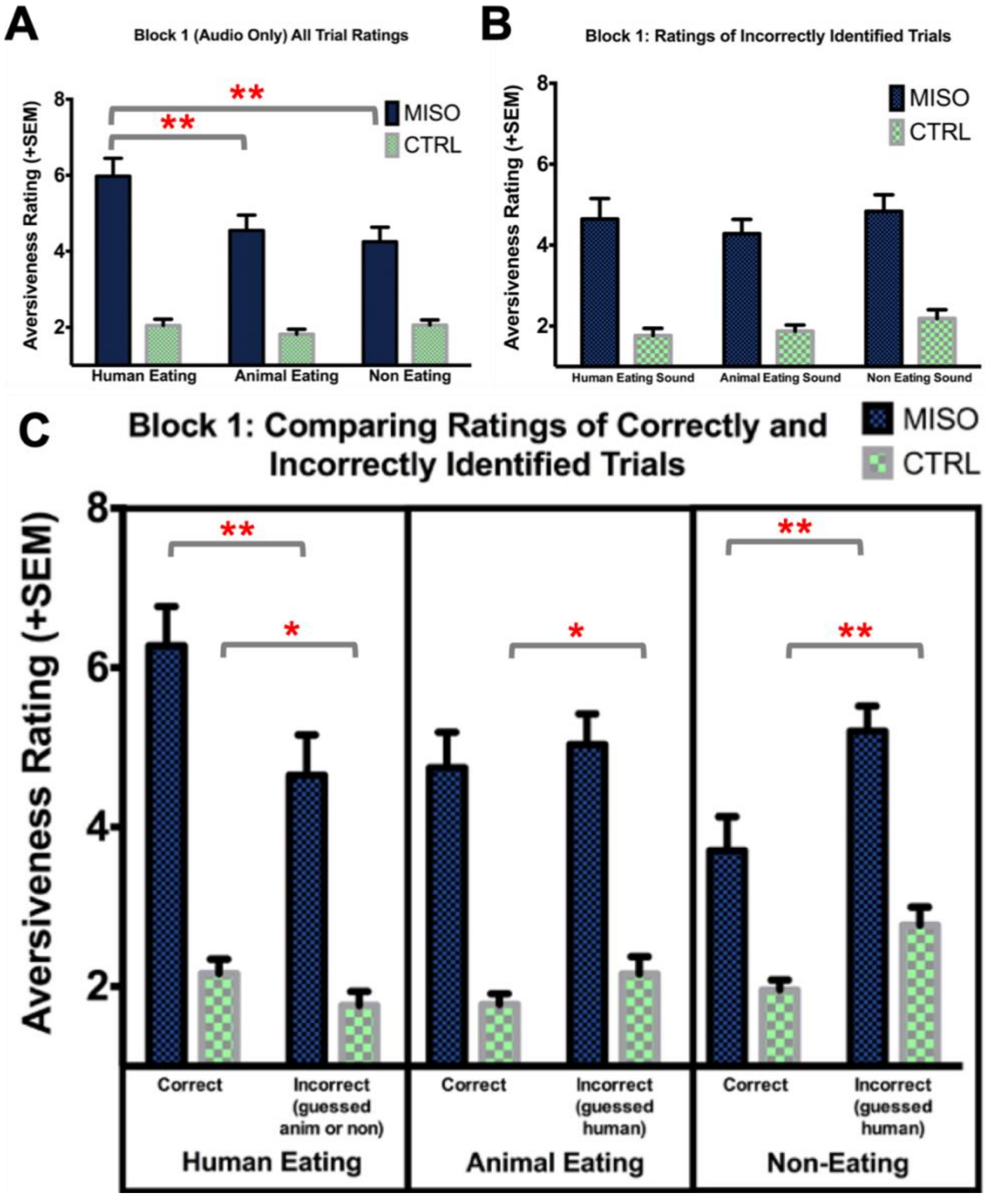

Figure 2.2. Block 1 aversiveness ratings. A) Average aversiveness ratings of human eating, animal eating and non eating sounds for misophonic and control participants in block 1, regardless of if the sound category was correctly identified. B) Average aversiveness ratings of incorrectly identified human eating, animal eating and non eating sounds for misophonic and control participants in block 1. C) Average aversiveness ratings of correctly and specific incorrectly identified human eating, animal eating and non eating sounds of misophonic and control participants in block 1. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

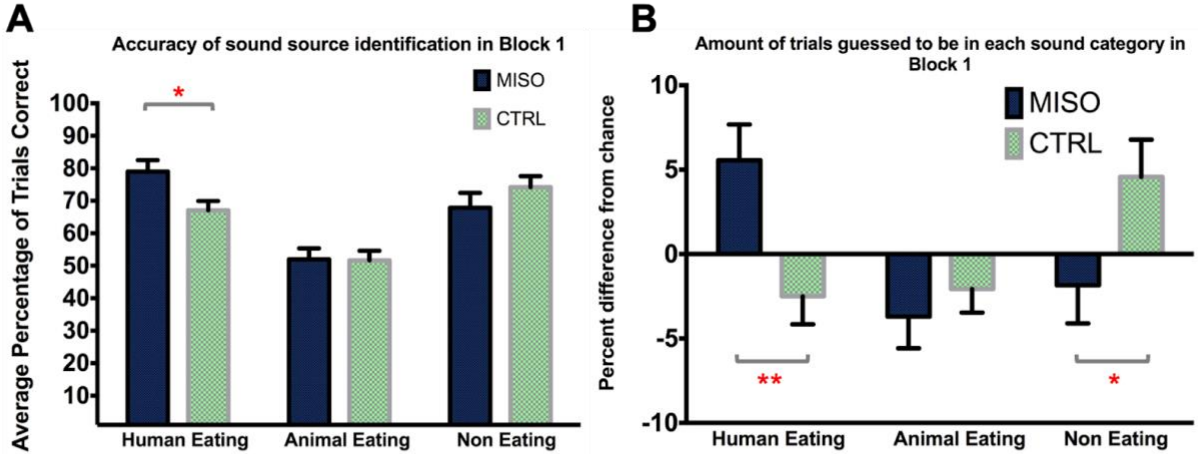

Figure 2.3. Sound category guess accuracy and propensity. A) Average percentage of correct human eating, animal eating and non eating trials of misophonic and control participants in block 1. B) Depicts how frequently trials (shown as percent difference from chance (33.33%)) were guessed by misophonic and control participants to be human eating, animal eating and non eating sounds in block 1. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

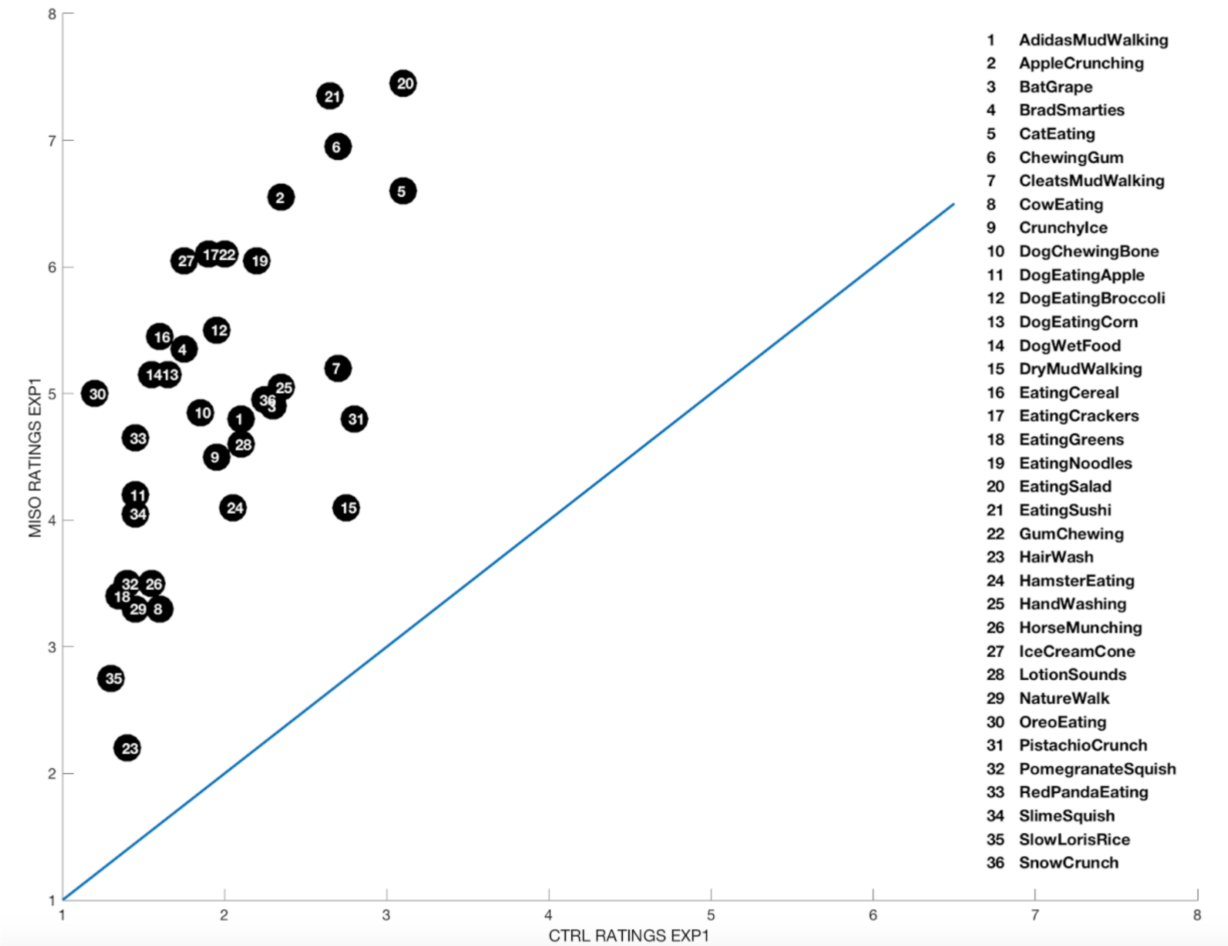

Figure 2.4. Scatterplot of misophonic and control ratings of all stimuli in block 1.

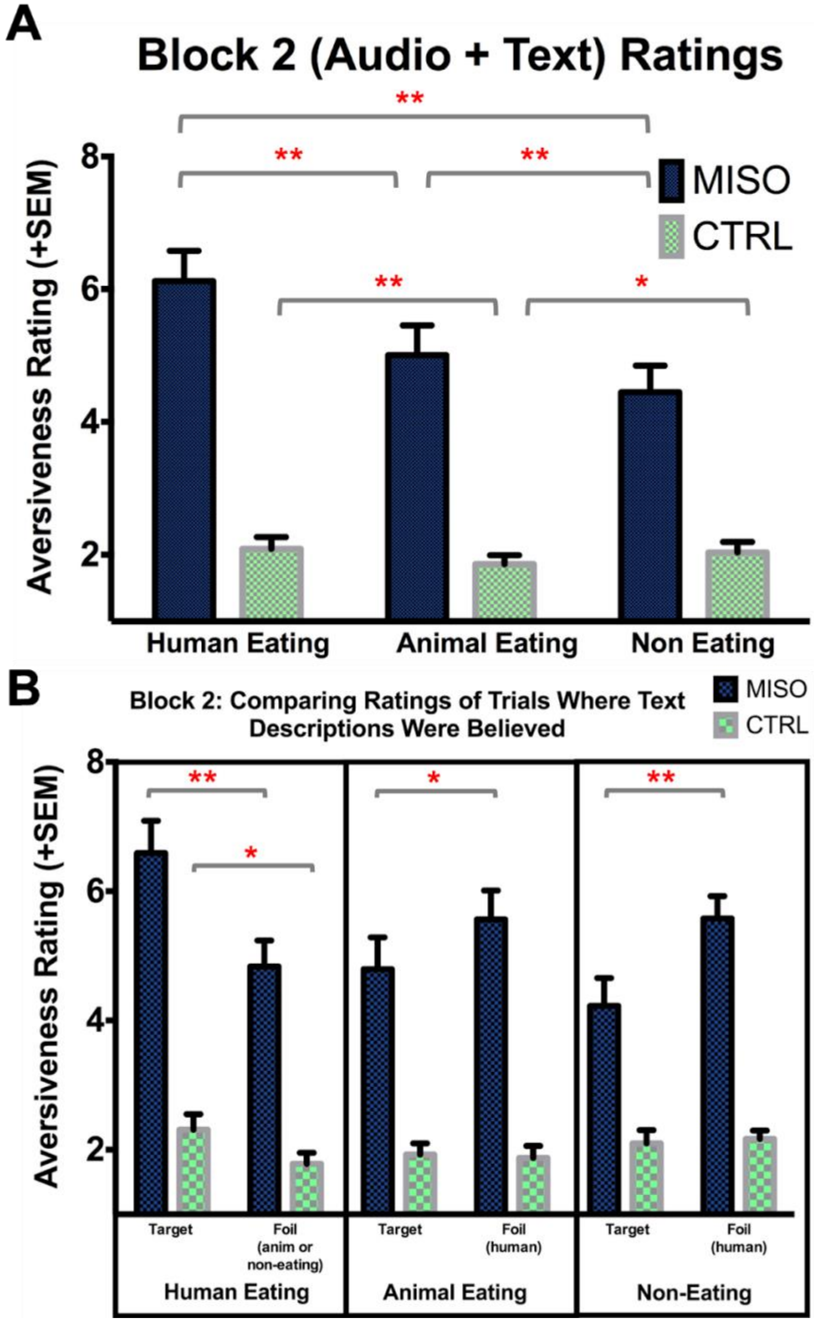

Figure 2.5. Block 2 aversiveness ratings. A) Average aversiveness ratings of human eating, animal eating and non eating sounds for misophonic and control participants in block 2, regardless of whether participants received an accurate (target) or false (foil) textual description and whether the sound category was correctly identified. B) Misophonic and control participants average aversiveness ratings of human eating, animal eating and non eating sounds that were either correctly identified as their target description (left) or incorrectly identified as specific foil descriptions (right) in block 2. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

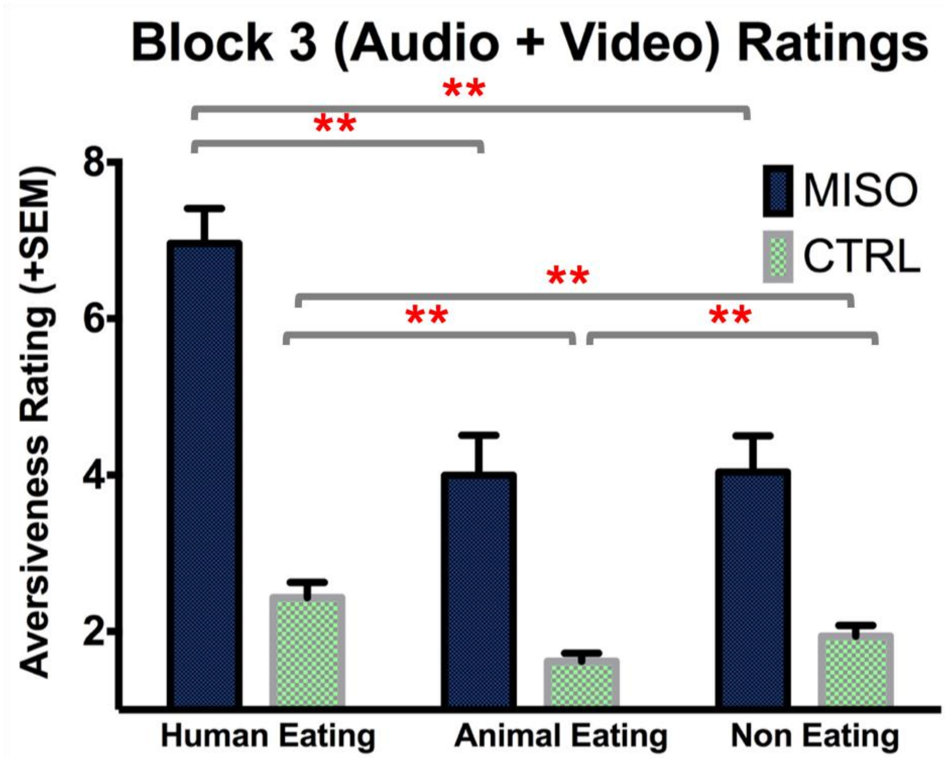

Figure 2.6. Block 3 aversiveness ratings. Average aversiveness ratings of human eating, animal eating and non eating sounds for misophonic and control participants in block 3. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

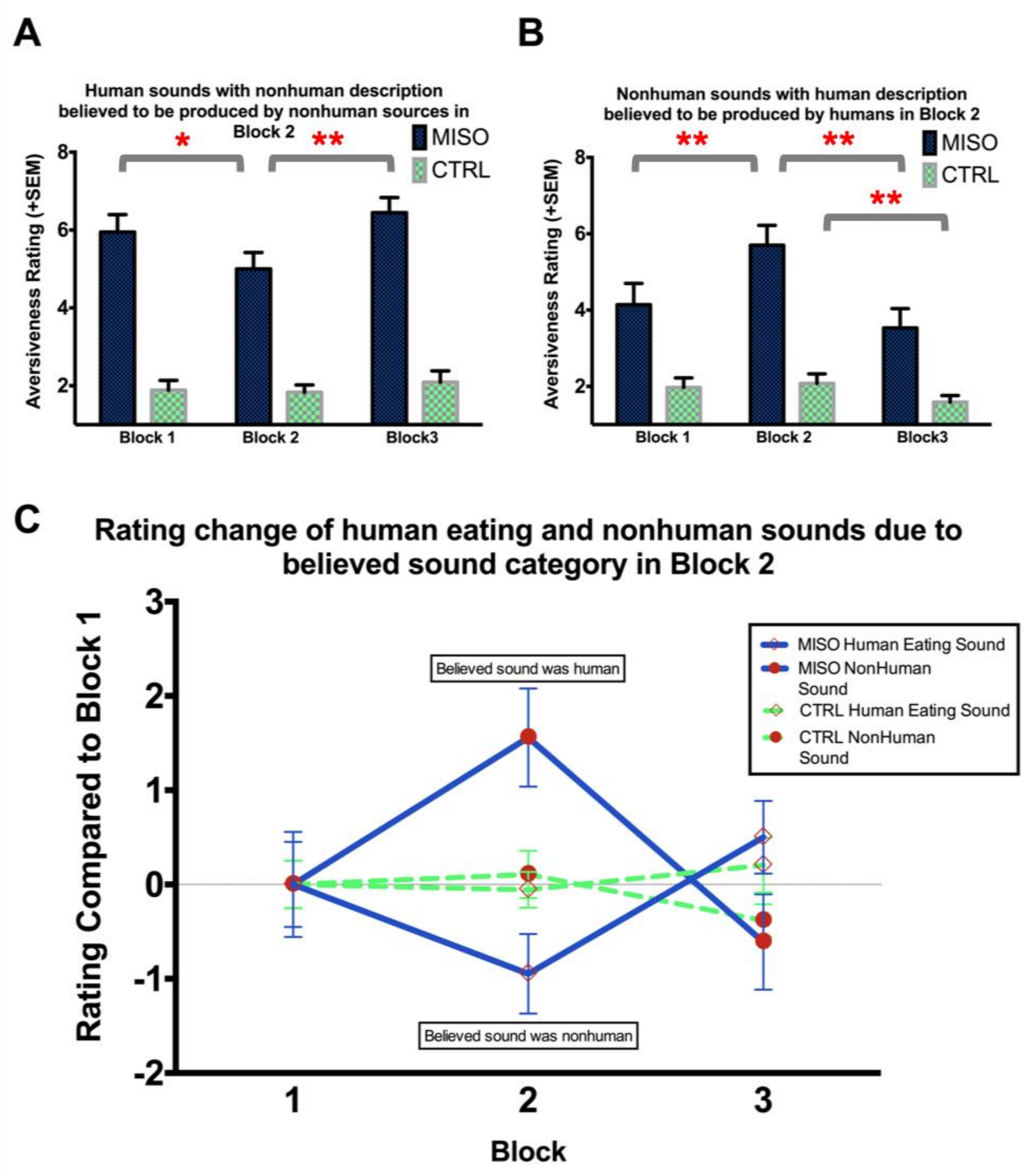

Figure 2.7. Rating change of human eating sounds and nonhuman sounds (believed to be the opposite type of sound in block 2) across blocks. A) Depicts the average misophonic and control aversiveness ratings of human eating sounds that were correctly identified as human eating sounds in block 1 but were believed to be produced by nonhuman sources in block 2. B) Depicts the average misophonic and control aversiveness ratings of nonhuman sounds that were correctly identified as nonhuman sounds in block 1 but were believed to be human eating sounds in block 2. C) Combines figures 7A and 7B into one graph but instead of displaying average aversiveness ratings, displays the average change in rating of each sound type in each block relative to block 1.

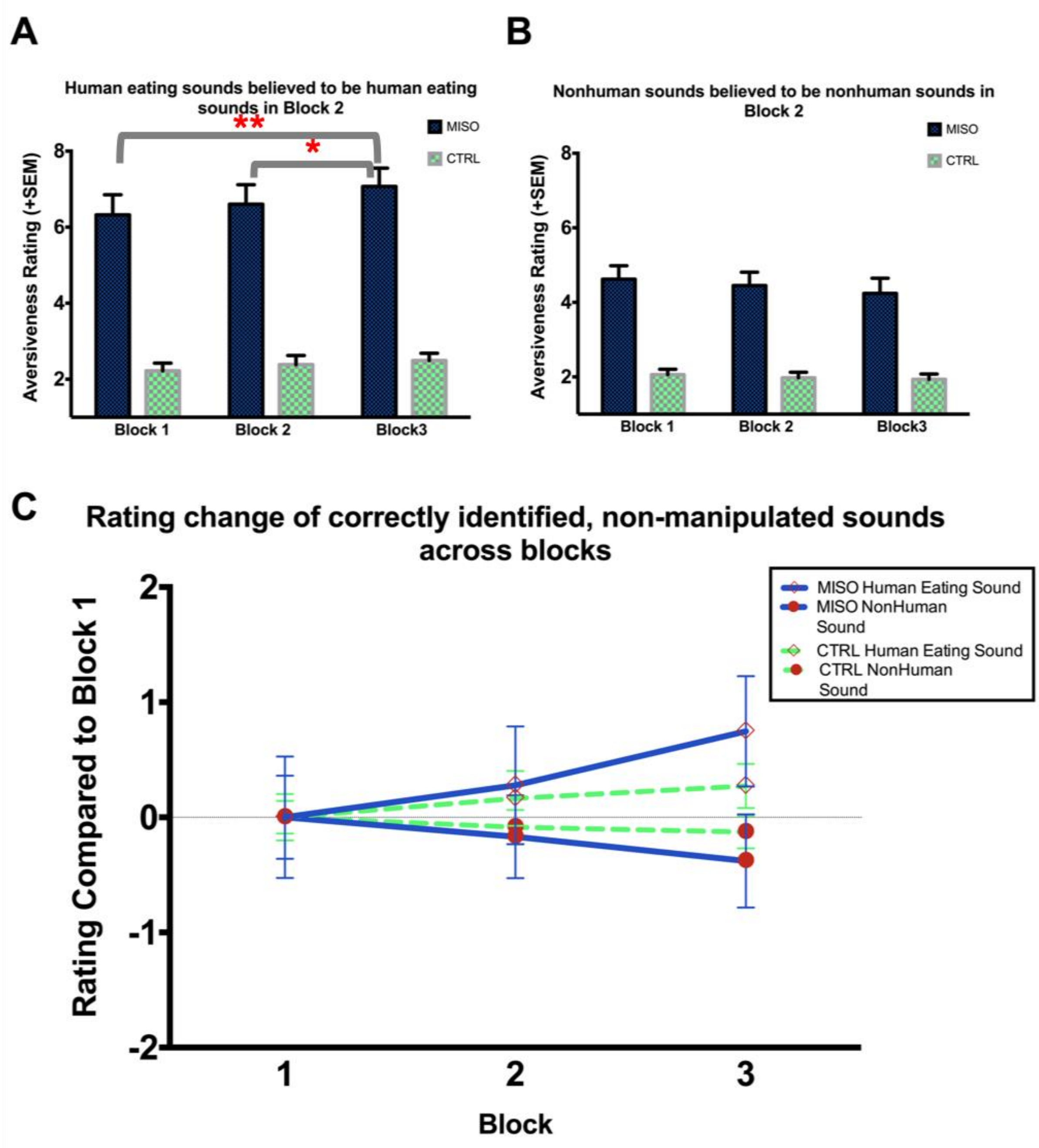

Figure 2.8. Rating change of human eating sounds and nonhuman sounds (that were correctly identified in all blocks) across blocks. A) Depicts the average misophonic and control aversiveness ratings of human eating sounds that were correctly identified as human eating sounds in blocks 1 and 2. B) Depicts the average misophonic and control aversiveness ratings of nonhuman sounds that were correctly identified as nonhuman sounds in blocks 1 and 2. C) Combines figures 8A and 8B into one graph but instead of displaying average aversiveness ratings, displays the average change in rating of each sound type in each block relative to block 1.

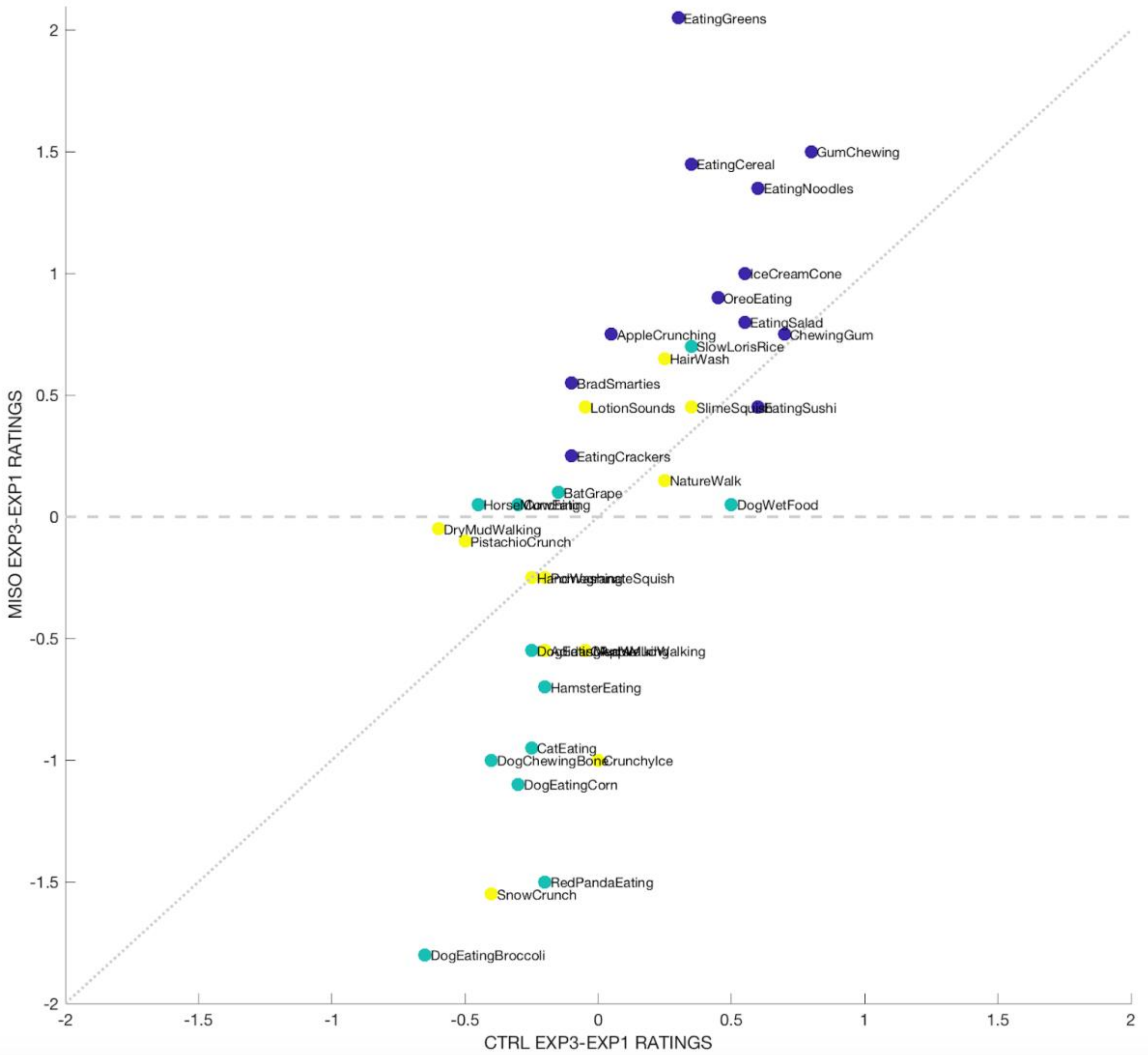

Figure 2.9. Rating change of individual stimuli between blocks 3 and 1. Shows the difference in rating of each stimulus from when it was encountered in block 1 (audio only) and block 3 (audio + video). Control rating differences are shown on the x axis and misophonic rating differences are shown on the y axis. Purple dots represent human eating sounds, turquoise dots represent animal eating sounds and yellow dots represent non eating sounds.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Edelstein, Miren; Monk, Bradley; Rouw, Romke; Ramachandran, V.S. The dissertation author was the primary investigator and author of this paper.

CHAPTER 3

Timbral expertise influences pitch labeling performance in absolute pitch possessors

Edelstein, M., Monk, B., Henthorn, T., and Deutsch, D.

ABSTRACT

Absolute pitch (AP), commonly known as perfect pitch, is the rare ability to identify auditory tones (typically as musical pitches) in isolation, without using a reference pitch. However, despite being widely studied, a strict definition as to what constitutes 'having absolute pitch' is lacking. Previous research suggests that not everyone who has AP abilities performs the same across different assessments, which suggests individuals may be using a variety of capacities or strategies to convert raw frequencies into note names. Indeed, note identification seems to be a complex process that takes into account multiple contextual factors, including timbre, note range, and note color. Our current study specifically investigates the interaction of sound properties (i.e. timbre, frequency, etc.) and personal experience (i.e. years of musical training, expertise with specific instruments, etc.) on pitch labeling ability.

Results support our hypothesis that the timbre that pitches are presented in can influence AP possessors' ability to label said pitches. Specifically, an AP possessor's accuracy and speed of pitch labeling appeared to improve when pitches were presented in the timbre of the instrument that they had the most expertise in. Conversely, speed and accuracy of pitch labeling was often diminished when pitches were presented in the timbre of non primary instruments. In general, pitch labeling ability was mediated by the amount of experience a musician had with the specific timbre being presented to them. This suggests there are more dimensions to AP than just the simple derivation of note names from raw frequencies. The finding that timbre can facilitate or hinder pitch labeling as a function of expertise, indicates that there are contextual nuances involved with AP that should be recognized and considered when formally assessing the ability.

INTRODUCTION

Absolute pitch (AP), also known as perfect pitch, is the rare ability to identify or produce a musical pitch without the aid of a reference pitch. The fact that AP is so uncommon even amongst serious, lifelong musicians, has piqued the curiosity of researchers for many decades and as a result, extensive research spanning the fields of psychology, neuroscience, linguistics and even genetics, has been conducted on the topic.

Many research studies have focused on addressing the question of why certain individuals but not others end up developing AP. One notable finding is that the vast majority of AP possessors receive musical training early on, usually between the ages of 4 & 6, suggesting that there may be a critical period associated with AP acquisition (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998; Deutsch, 2013; Gregersen, Kowalsky, Kohn, & Marvin, 1999, 2001; Levitin & Zatorre, 2003; Takeuchi & Hulse, 1993). However, the fact that many experienced musicians who receive early musical training do not go on to develop AP indicates that the critical period hypothesis might not provide the full picture and that certain genes may play a role in AP acquisition as well (Baharloo et al., 1998; Baharloo, Service, Risch, Gitschier, & Freimer, 2000; Gregersen et al., 1999, 2001; Theusch, Basu, & Gitschier, 2009; Theusch & Gitschier, 2011). There is also evidence that AP possession is far more prevalent in musicians who speak a tone language, such as Mandarin Chinese or Vietnamese, than in musicians who speak non-tone languages, such as English (Deutsch, Li, & Shen, 2013; Deutsch, 2006; Deutsch, Le, et al., 2009; Lee & Lee, 2010). As the inflection of a pitch can change the semantic meaning of a word in tone languages, it makes sense that these individuals, who have learned from a young age to form associations between pitches and verbal labels, may be predisposed to developing AP later in life. However, despite the strong link between speaking a tone language and AP, there are indeed musicians who speak non-tone languages who end up acquiring AP, although they are few and far between. Because a very small percentage of these non-tone language speaking musicians develop AP and the majority do not (even with equivalent years and onset of musical training) it suggests that the ones who do may have different underlying neurological or cognitive mechanisms. Indeed, a study conducted by Deutsch & Dooley (2013) found that non-tone language speaking AP possessors had a significantly larger auditory digit span than a control group of non-tone language speaking AP nonpossessors with equivalent musical experience. Additionally, imaging studies have identified significant structural and functional differences in the brains of musicians with and without AP and shown that many brain regions associated with AP (namely temporal and frontal areas) are also known to be associated with categorization, language, speech and pitch processing (Keenan, Thangaraj, Halpern, & Schlaug, 2001; Loui, Li, Hohmann, & Schlaug, 2011; Oechslin, Meyer, & Jäncke, 2010; Ohnishi et al., 2001; Schlaug, Jäncke, Huang, & Steinmetz, 1995; Wengenroth et al., 2013; Zatorre, 2003).

Other research studies have focused less on identifying the potential underlying causes of AP and more on addressing the challenge of characterizing the multidimensional nature of the ability. Research suggests that AP is not as straightforward as simply converting raw frequencies into note names and is instead a much more nuanced process that can be modulated by several factors, such as timbre, pitch range and note color (Bahr, Christensen, & Bahr, 2005; Brammer, 1951; Levitin & Rogers, 2005; Lockhead & Byrd, 1981; Marvin & Brinkman, 2000; Miyazaki, 1988, 1989, 1990; Takeuchi and Hulse, 1993; Vanzella & Schellenberg, 2010; Ward, 1999; Wong & Wong, 2014). In the case of timbre, it has been shown that certain ones tend to be much more accessible than others when it comes to pitch identification. Namely, it is generally much easier to label a pitch when it is presented as a piano tone than when it is presented as a sine wave tone (Athos et al., 2007; Baharloo et al., 1998; Bahr, Christensen, & Bahr, 2005; Deutsch, 2013; Lee & Lee, 2010; Lockhead & Byrd, 1981; Miyazaki, 1989; Takeuchi & Hulse, 1993; Vanzella & Schellenberg, 2010; Wong & Wong, 2014). This is likely due to how often certain timbres are encountered in the world. Because the average musician probably doesn't come across sine wave tones as frequently as piano tones, they will have had fewer total exposures and therefore less experience with them, which would explain why it is more difficult to label a pitch when it is presented as a sine wave tone.

However, although there have been quite a few research studies confirming general effects of timbre on AP performance, interestingly enough, the majority of these studies do not take into account the specific timbral expertise of their AP participants. Since the effect of timbre on AP performance appears to be driven by level of familiarity one has with that timbre (Sergeant, 1969), this suggests that AP possessors should perform the best on a pitch labeling task when the pitches presented are in the timbre of the instrument they have the most expertise in. The current study tested this idea by extending previous work on timbre and AP, while also making several new examinations.

The study compared the performance of long-term pianists and violinists (who self-identified as having AP) on AP tests that were given in both piano and violin timbre. In order to examine the effect of expertise, participants were required to identify as being primarily a pianist or a violinist but not both. We hypothesized that a congruency effect between instrument played and note timbre would be observed such that pianists would be more accurate and faster at identifying piano tones than violin tones, and violinists would be more accurate and faster at identifying violin tones than piano tones. In other words, we expected a participant to perform better when the timbre of the AP test matched the timbre of their instrument of expertise than when it did not. By selecting two groups of AP possessors who specifically had expertise in one of two popular instruments and testing both groups on notes played on both instruments, the current study is able to isolate and examine the effect of specific timbral expertise on AP performance in a way that most previous studies have not.

METHODS

Participants

Twenty-four individuals with normal hearing from the UCSD student body and southern California area who self-identified themselves as having AP participated in the study. Thirteen were pianists (8 males; average age = 25.23, range = 18-35) and eleven were violinists (2 males; average age = 29, range = 18-42). Pianists began piano lessons at 4.58 years of age on average and continued piano lessons for an average of 14.58 years. Violinists began violin lessons at 5.86 years of age on average and continued violin lessons for an average of 14.95 years. Although participants were allowed to have played or currently play multiple instruments, the main requirement was that their long term (and current), primary instrument was either piano or violin (but not both). In addition to violin and piano, participants reported having also played the clarinet, viola, cello, drums, flute and voice (Fig. 1A).

Materials & Procedure

Participants were presented with two experimental blocks with 48 trials in each. During each trial, participants heard a tone between C4 (middle C) and B5 (a two-octave range). Each individual tone was repeated twice per block for a total of 48 trials per block and 96 trials for the entire experiment. The tones in one of the experimental blocks consisted of violin tones while the tones in the other consisted of piano tones. In order to familiarize participants with the experimental task, five practice trials were given prior to each experimental block, which consisted of tones in the same timbre as the block they preceded.

All tones were generated in Logic Pro, tuned to A = 440Hz and presented to participants on a MacBook Pro in MATLAB, through a pair of Sennheiser headsets at a comfortable volume. Each tone was played for 500ms followed by 4.25 seconds of silence before the next tone began. Participants were instructed to identify each tone as quickly and accurately as possible by pressing a button with the correct note label on a Korg nanoPAD2 keypad. The keypad had twelve buttons, each of which corresponded with one of the twelve Western pitch classes. In addition to accuracy, these button presses also registered reaction times, which were recorded in seconds.

Tones were presented in two semi-random orders where each successive tone was at least 4 semitones apart from the previous tone, meant to minimize the use of relative pitch cues when making pitch judgments. This tone range (C4 to B5) was selected because it is well within the range of both violin and piano and was intended to mitigate potentially confounding effects of instrumental pitch range (Miyazaki, 1989). In order to account for potential ordering effects, the two experimental blocks were counterbalanced by both tone and timbre order (Fig. 1B).

Scoring

The study had three dependent variables: number of correct trials, reaction time and the number of semitones deviated from the correct answer. Trials where participants were off by one semitone were not counted as correct. A scoring technique similar to one described by Bermudez

& Zatorre (2009) was utilized to determine the number of semitones participants deviated from the correct answer.

RESULTS

Descriptive:

Although all participants self-identified as having AP prior to participating in the study, performance on the experimental task (which simultaneously served as an AP screening test) ranged widely (100% to 21.8% of trials correct when averaged across both timbre conditions, not allowing for semitone errors). Despite the large range, all participants performed well above chance (which was 8.3% of trials correct). Figure 1D depicts the distribution of pianists and violinists who fell into various performance quartiles on the experimental task. Overall, the majority of participants (15 out of 24) scored well enough to be placed in the top quartile (75% and above).

Correlations:

A significant negative correlation between average reaction time and average pitch labeling accuracy was found [$r = -.87$, $n = 24$, $p < .001$] indicating that in general, the more accurate a participant's pitch labeling abilities, the faster their response (Fig. 1C).

A significant negative correlation between reaction time and trial correctness was found [$r = -.507$, $n = 2304$, $p < .001$], showing that participants tended to respond faster to trials that they got correct and slower to trials that they got incorrect. This indicates that participants likely experienced more uncertainty about trials they ended up getting wrong, which shows up in the form of slower reaction times.

Neither years of training on one's primary instrument nor age of onset of learning one's primary instrument significantly correlated with number of correct trials, semitone accuracy or reaction time.

Between subjects factor (primary instrument: pianists vs violinists)

Correct Trials: When considering responses to all trials (of both piano and violin timbre) by all participants, participants got 78.7% of trials correct (M = 75.63 (out of 96) trials, SD = 24.69). Pianists got 81.2% of trials correct (M = 77.92 (out of 96) trials, SD = 25.73) and violinists got 75.9% of trials correct (M = 72.91 (out of 96) trials, SD = 24.35), but this difference was not significant, [$p > .05$].

Semitone Accuracy: When considering responses to all trials (both correct and incorrect) by all participants, participants were .416 semitones (SD = .624) off from guessing the correct note on average. Pianists (M = .4 semitones off, SD = .7) and violinists (M = .44 semitones off, SD = .505) did not differ significantly when it came to overall semitone accuracy, [$p > .05$]. When only considering incorrect trials, participants were 1.2 semitones (SD = .833) off on average. Pianists (M = 1.06 semitones off, SD = .82) and violinists (M = 1.37 semitones off, SD = .856) did not differ significantly on semitone accuracy on incorrect trials, [$p > .05$].

Reaction Time: When considering reaction time of all trials (of both piano and violin timbre) by all participants, participants responded in an average of 1.81 seconds (SD = .586). Pianists (M = 1.65 seconds, SD = .6) and violinists (M = 1.96 seconds, SD = .551) did not differ significantly on reaction time overall, [$p > .05$]. When only considering incorrect trials, participants responded in 2.233 seconds (SD = .616) on average. Interestingly, pianists (M = 1.98

seconds, SD = .65) responded significantly faster than violinists (M = 2.57 seconds, SD = .39) on incorrect trials [t(19) = -2.382, p < .05].

Within subjects factor (trial type: piano vs violin timbre)

Correct Trials: On average, participants got 78.9% of piano trials correct (M = 37.88 (out of 48) trials, SD = 12.26) and 78.6% of violin trials correct (M = 37.75 (out of 48) trials, SD = 12.955) overall. In terms of number of correct trials, overall performance on piano and violin trials did not differ significantly, [p > .05].

Semitone Accuracy: When averaging across both correct and incorrect responses, participants were .414 semitones (SD = .586) off on piano trials and .419 semitones (SD = .673) off on violin trials, on average. This difference was not significant, [p > .05]. When only considering incorrect trials, participants were off by 1.24 semitones on piano trials (SD = 1.09) and 1.17 semitones on violin trials (SD = .82) but this difference was not significant, [p > .05].

Reaction Time: When considering reaction time, on average, participants responded in 1.78 seconds (SD = .578) to piano trials and 1.81 seconds (SD = .617) to violin trials. In terms of reaction time, piano and violin trials did not differ significantly, [p > .05].

Effect of instrumental expertise: **primary instrument x trial type**

By having participants with expertise in either piano or violin respond to trials in both piano and violin timbres, we were able to investigate the potential effect of specific timbral expertise on AP performance, which was the primary aim of this study. Our hypothesis stated that participants should demonstrate better AP performance when trials are in the timbre of their instrument of expertise as opposed to the timbre of another instrument. Our findings strongly supported this hypothesis.

Correct Trials: Participants got significantly more trials correct when the timbre of the trials was congruent with the participant's primary instrument (M = 39.13 (out of 48) trials, SD = 11.26) as opposed to incongruent with it (M = 36.5 (out of 48) trials, SD = 13.7), [$t(23) = 2.935$, $p < .01$] (Fig. 2A-D). Specifically, pianists got significantly more trials correct when identifying piano tones (M = 40.23 (out of 48) trials, SD = 11.39) than violin tones (M = 37.69 (out of 48) trials, SD = 14.53) [$t(12) = 2.05$, $p < .05$] and violinists got significantly more trials correct when identifying violin tones (M = 37.82 (out of 48) trials, SD = 11.51) than piano tones (M = 35.09 (out of 48) trials, SD = 13.19) [$t(10) = 2.012$, $p < .05$] (Fig. 2E).

Although pianists got more piano trials correct than violinists did, this difference was not significant, [$p > .05$]. Additionally, violinists did not get significantly more violin trials correct than pianists did, [$p > .05$].

Semitone Accuracy: When considering both correct and incorrect trials, participants were off by fewer semitones when identifying pitches if the timbre of the trial was congruent with the participant's primary instrument (M = .329 semitones off, SD = .49) as opposed to incongruent with it (M = .504 semitones off, SD = .733), [$t(23) = -2.99$, $p < .01$] (Fig. 3A-D). Specifically, pianists were off by significantly fewer semitones when identifying piano tones (M = .31 semitones off, SD = .57) than violin tones (M = .48 semitones off, SD = .85) [$t(12) = -1.865$, $p < .05$] and violinists were off by significantly fewer semitones when identifying violin tones (M = .35 semitones off, SD = .41) than piano tones (M = .53 semitones off, SD = .61) [$t(10) = -2.407$, $p < .05$] (Fig. 3E). For a descriptive illustration of semitone accuracy in a small subset of participants, please see Figure 5.

Although pianists were off by fewer semitones than violinists on piano trials, this difference was not significant, [$p > .05$], and although violinists were off by fewer semitones than pianists on violin trials, this difference was not significant either, [$p > .05$].

When only considering incorrect trials, participants were off by fewer semitones when identifying pitches if the timbre of the trial was congruent with the participant's primary instrument (M = .93 semitones off, SD = .85), as opposed to incongruent with it (M = 1.48, SD = .96), [($t(23) = -3.338$, $p < .01$]. Pianists were off by significantly fewer semitones when identifying piano tones (M = .84, SD = .92) than violin tones (M = 1.28, SD = .84) [$t(12) = -2.448$, $p < .05$] and violinists were off by significantly fewer semitones when identifying violin tones (M = 1.03 semitones off, SD = .83) than piano tones (M = 1.72, SD = 1.13) [$t(10) = -2.302$, $p < .05$].

Reaction Time: Participants were significantly faster at identifying pitches if the timbre of that trial was congruent with the participant's primary instrument (M = 1.76 seconds, SD = .58) as opposed to incongruent with it (M = 1.84 seconds, SD = .61), [$t(23) = -1.754$, $p < .05$] (Fig. 4A-D). Specifically, pianists were significantly faster at identifying piano tones (M = 1.6 seconds, SD = .56) than violin tones (M = 1.71 seconds, SD = .65) [$t(12) = -2.034$, $p < .05$] but violinists were not found to be significantly faster at identifying violin tones (M = 1.94 seconds, SD = .58) than piano tones (M = 1.99 seconds, SD = .55) [$p > .05$] (Fig. 4E).

Pianists were significantly faster than violinists on piano trials, [$t(22) = -1.697$, $p = .05$] and interestingly enough, were also faster than violinists on violin trials, although not to a significant degree [$p > .05$].

DISCUSSION

Results strongly supported our hypothesis that AP performance is improved when pitches are presented in the timbre of one's instrument of expertise as opposed to another timbre (even if both of those timbres are of commonly encountered instruments). Although previous studies have investigated how AP possessors perform better overall with some timbres and worse with others, most do not take into consideration the instrumental expertise of their participants. Piano is frequently cited as one of the easiest timbres to identify pitches in (likely due to its ubiquity in Western music and the fact that many musicians have received piano training at some point during their career) and while this may be true on average, the findings from this study suggest that AP possessors will likely still perform better in the timbre of one's primary instrument. By evaluating participants who were either long term pianists or violinists and administering AP tests in both piano and violin timbre to each group, this study was able to quantitatively assess the influence of a given participant's instrumental expertise as a performance determinant.

When comparing a participant's performance on trials where the timbre matched their primary instrument (expertise) and trials where the timbre didn't match their primary instrument (non expertise), we found that they performed significantly better on these expertise trials than on non expertise trials in all three areas of assessment (number of correct trials, average semitone error distance and reaction time). In other words, pianists performed significantly better on trials presented in piano tones than trials presented in violin tones and violinists performed significantly better on trials presented in violin tones than trials presented in piano tones.

The findings from this study are relevant as the vast majority of AP screening tests administered for research purposes use only piano or pure tones as test tones, making it very possible that many AP possessors are not being properly assessed for AP ability. As our results show that expertise with a timbre can impact several measures of AP performance (including number of correct trials, reaction time and semitone accuracy) to a statistically significant degree, researchers should consider supplementing current AP screening tests with ones that include the timbre of the participant's primary instrument in order to more accurately characterize a participant's AP.

One of the biggest limitations of the current study is its small sample size, which unfortunately is a common issue for research studies that investigate rare populations such as individuals with AP. However, it is possible that with a larger sample, certain consistent and intriguing trends that did not emerge as significant in this study might become more pronounced. In particular, although no main effect of primary instrument was found, pianists did tend to perform better than violinists overall in terms of number of correct trials, semitone accuracy and reaction time (just not to a statistically significant degree). It would be interesting to investigate this trend further and see if it holds up with timbres that participants are less familiar with or ones that are harmonically dissimilar to piano and violin. Additionally, supplemental follow up studies investigating if AP performance is influenced by how frequently, recently and continuously an individual has played their primary instrument would also be worthwhile.

References

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(37), 14795–800. doi:10.1073/pnas.0703868104

Baharloo, S., Johnston, P. a, Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute pitch: an approach for identification of genetic and nongenetic components. *American Journal of Human Genetics*, *62*(2), 224–31. doi:10.1086/301704

Baharloo, S., Service, S. K., Risch, N., Gitschier, J., & Freimer, N. B. (2000). Familial aggregation of absolute pitch. *American Journal of Human Genetics*, *67*(3), 755–8. doi:10.1086/303057

Bahr, N., Christensen, C. A., & Bahr, M. (2005). Diversity of accuracy profiles for absolute pitch recognition. *Psychology of Music*, *33*(1), 58–93. doi:10.1177/0305735605048014

Bermudez, P., & Zatorre, R. J. (2009). A distribution of absolute pitch ability as revealed by computerized testing. *Music Perception*, 27(2), 89-101.

Brammer, L. M. (1951). Sensory cues in pitch judgment. *Journal of experimental psychology*, 41(5), 336.

Deutsch, D. (2006). The Enigma of Absolute Pitch. *Acoustics Today*, *2*(4).

Deutsch, D. (2013). Absolute Pitch. In D. Deutsch (Ed.), *The Psychology Of Music* (3rd Edition., pp. 141–182). San Diego: Elsevier Inc

Deutsch, D., & Dooley, K. (2013). Absolute pitch is associated with a large auditory digit span: a clue to its genesis. *The Journal of the Acoustical Society of America*, *133*(4), 1859–61. doi:10.1121/1.4792217

Deutsch, D., Le, J., Dooley, K., Henthorn, T., Shen, J., & Head, B. (2009). Absolute Pitch and Tone Language: Two New Studies. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, (Escom), 69–73.

Deutsch, D., Li, X., & Shen, J. (2013). Absolute pitch among students at the Shanghai Conservatory of Music: a large-scale direct-test study. *The Journal of the Acoustical Society of America*, *134*(5), 3853–9. doi:10.1121/1.4824450

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (1999). Absolute pitch: prevalence, ethnic variation, and estimation of the genetic component. *American Journal of Human Genetics*, *65*(2), 911–913. doi:10.2345/0899-8205-48.2.75

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (2001). Early Childhood Music Education and Predisposition to Absolute Pitch: Teasing Apart Genes and Environment. *American Journal of Medical Genetics*, *98*(2000), 280–282.

Keenan, J. P., Thangaraj, V., Halpern, A. R., & Schlaug, G. (2001). Absolute pitch and planum temporale. *NeuroImage*, *14*(6), 1402–8. doi:10.1006/nimg.2001.0925

Lee, C.-Y., & Lee, Y.-F. (2010). Perception of musical pitch and lexical tones by Mandarin-speaking musicians. *The Journal of the Acoustical Society of America*, *127*(1), 481–90. doi:10.1121/1.3266683

Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends in Cognitive Sciences*, *9*(1), 26–33. doi:10.1016/j.tics.2004.11.007

Levitin, D. J., & Zatorre, R. J. (2003). On the Nature of Early Music Training and Absolute Pitch: A Reply to Brown, Sachs, Cammuso, and Folstein. *Music Perception*, *21*(1), 105–110.

Lockhead, G. R., & Byrd, R. (1981). Practically perfect pitch. *The Journal of the Acoustical Society of America*, 70(2), 387-389.

Loui, P., Li, H. C., Hohmann, A., & Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity. *Journal of Cognitive Neuroscience*, *23*(4), 1015–26. doi:10.1162/jocn.2010.21500

Marvin, E. W., & Brinkman, A. R. (2000). The effect of key color and timbre on absolute pitch recognition in musical contexts. *Music Perception: An Interdisciplinary Journal*, 18(2), 111-137.

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics*, *44*(6), 501–12.

Miyazaki, K. (1989). Absolute Pitch Identification: Effects of Timbre and Pitch Region. *Music Perception*, *7*(1), 1–14.

Miyazaki, K. (1990). The speed of musical pitch identification by absolute-pitch possessors. *Music Perception*, 8(2), 177-188.

Oechslin, M. S., Meyer, M., & Jäncke, L. (2010). Absolute pitch--functional evidence of speech-relevant auditory acuity. *Cerebral Cortex*, *20*(2), 447–55. doi:10.1093/cercor/bhp113

Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., & Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex*, *11*(8), 754–60.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*(5198), 699–701.

Sergeant, D. (1969). Experimental Investigation of Absolute Pitch. *Journal of Research in Music Education*, *17*(1), 135. doi:10.2307/3344200

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, *113*(2), 345–61.

Theusch, E., Basu, A., & Gitschier, J. (2009). Genome-wide Study of Families with Absolute Pitch Reveals Linkage to 8q24.21 and Locus Heterogeneity. *American Journal of Human Genetics*, *85*(1), 112–119.

Theusch, E., & Gitschier, J. (2011). Absolute pitch twin study and segregation analysis. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, *14*(2), 173–8. doi:10.1375/twin.14.2.173

Vanzella, P., & Schellenberg, E. G. (2010). Absolute pitch: Effects of timbre on note-naming ability. *PloS ONE*, 5(11), e15449.

Ward, W. D. (1999). Absolute pitch. In *The Psychology of Music* (Deutsch, D. ed.), pp. 265–298, Academic Press

Wengenroth, M., Blatow, M., Heinecke, A., Reinhardt, J., Stippich, C., Hofmann, E., & Schneider, P. (2013). Increased Volume and Function of Right Auditory Cortex as a Marker for Absolute Pitch. *Cerebral Cortex*, (May), 1127–1137. doi:10.1093/cercor/bhs391

Wong, Y. K., & Wong, A. C. N. (2014). Absolute pitch memory: Its prevalence among musicians and dependence on the testing context. *Psychonomic bulletin & review*, 21(2), 534-542.

Zatorre, R. J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience*, *6*(7), 692–5. doi:10.1038/nn1085

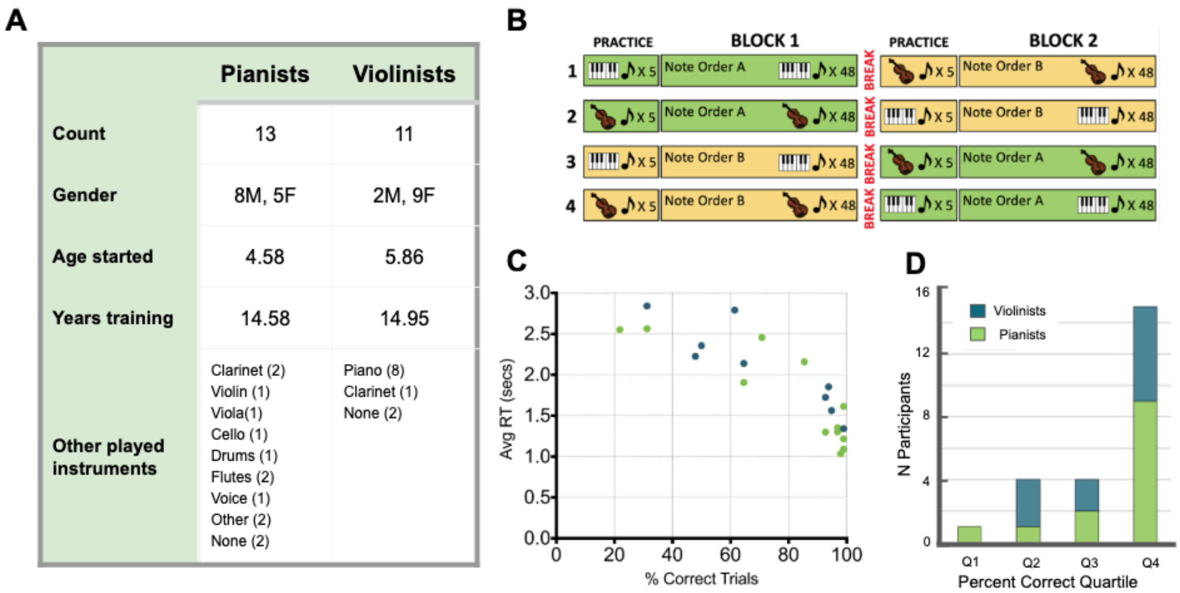

Figure 3.1. Descriptive and qualitative statistics. A) Describes the musical background of the study participants. B) Depicts the study's structure and four counterbalancing orders. Each participant was assigned to one of the four possible counterbalancing orders. C) Shows the significant correlation between the average percentage of correct trials and average reaction time of each participant. Violinists are depicted as the blue dots and pianists are depicted as the green dots. D) Histogram that shows the number of violinists (blue) and pianists (green) who fell into different performance quartiles (percentage of correct trials overall) on the experimental task.

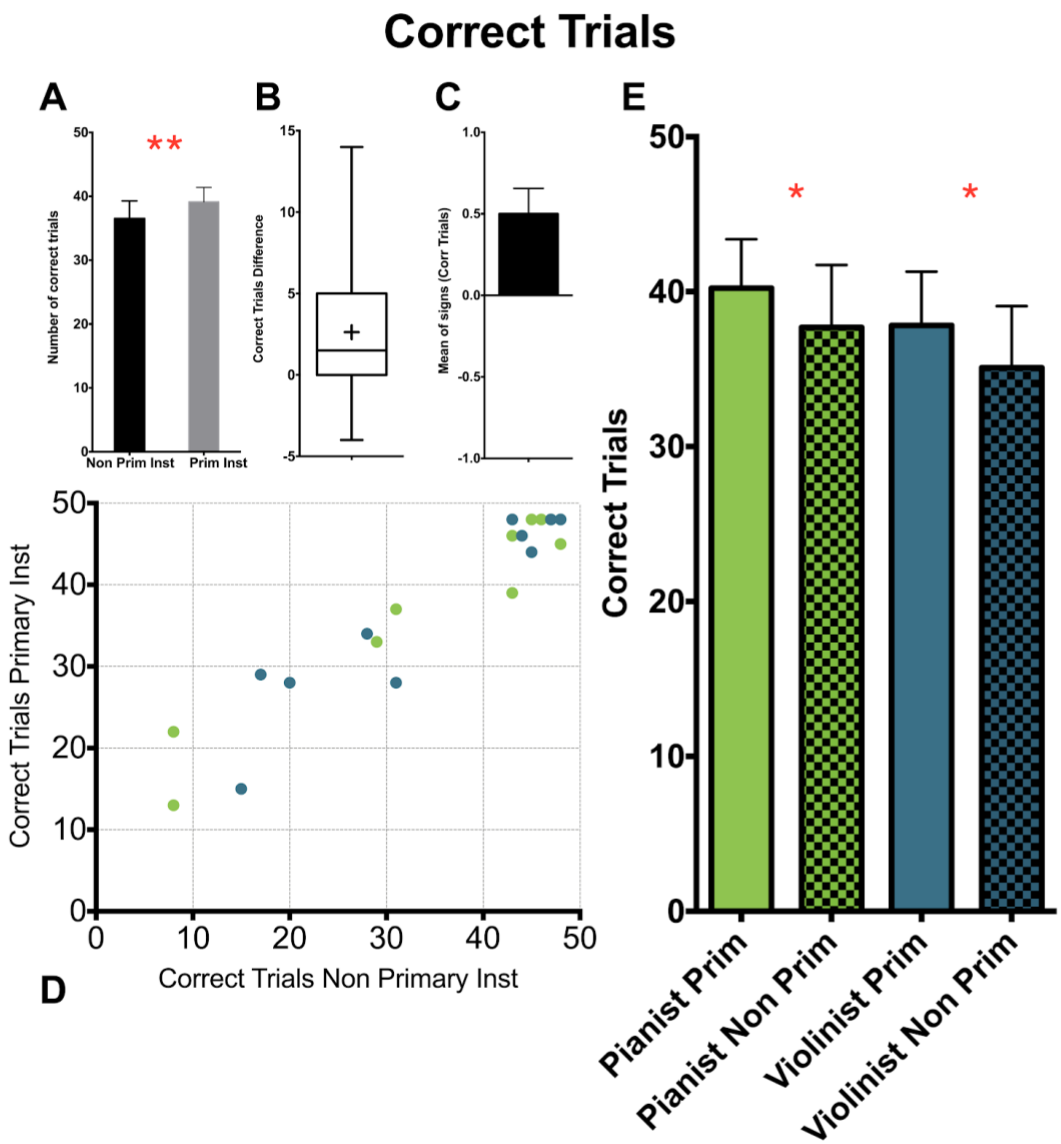

Figure 3.2. The effect of timbral expertise on number of correct trials. A) Depicts the average number of correct trials (out of 48) that were presented in the timbre of participants' primary and non primary instrument. B) Box plot depicting the average difference between the number of trials that participants got correct when the trial timbre was in their primary and non primary instrument. C) Depicts the mean of signs (as associated with a sign test) of instances where participants got more trials correct in the timbre of their primary instrument as opposed to their non primary instrument. D) A scatterplot depicting the number of trials each participant got correct (out of 48) in the timbre of their non primary instrument (x-axis) and primary instrument (y-axis). Pianists are represented by green dots and violinists are represented by blue dots. E) Depicts the number of trials that pianists and violinists got correct (out of 48) in the timbre of their primary and non primary instrument. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

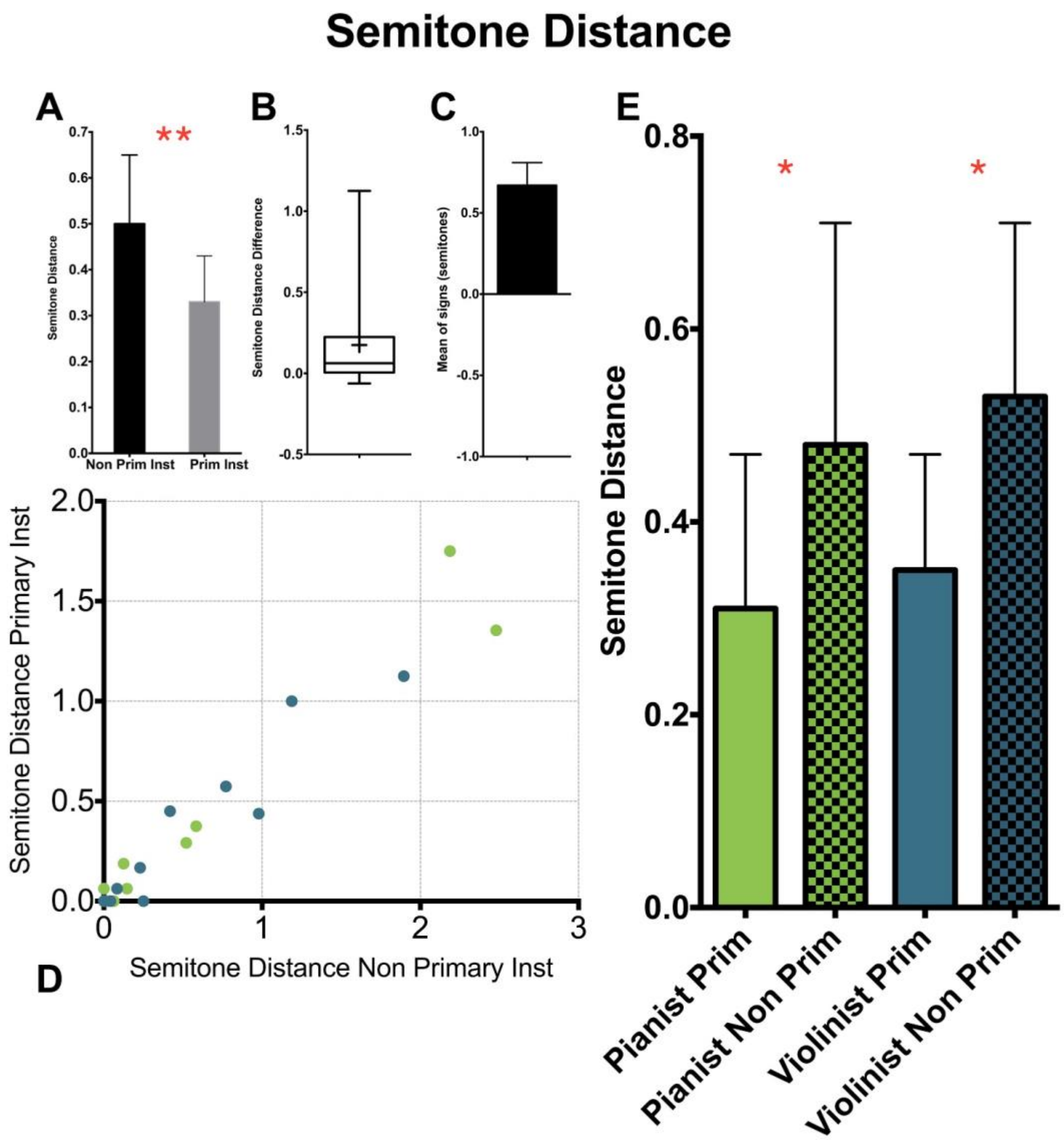

Figure 3.3. The effect of timbral expertise on semitone error distance. A) Depicts the average number of semitones that participants were off from the correct note by for trials in the timbre of their primary and non primary instrument. B) Box plot depicting the average difference between the number of semitones that participants were off by when the trial timbre was in their non primary and primary instrument. C) Depicts the mean of signs (as associated with a sign test) of instances where participants were off by more semitones when the trial timbre was in their non primary as opposed to their primary instrument. D) A scatterplot depicting the average number of semitones that participants were off by for trials in the timbre of their non primary instrument (x-axis) and primary instrument (y-axis). Pianists are represented by green dots and violinists are represented by blue dots. E) Depicts the average number of semitones that pianists and violinists were off by for trials in the timbre of their primary and non primary instrument. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

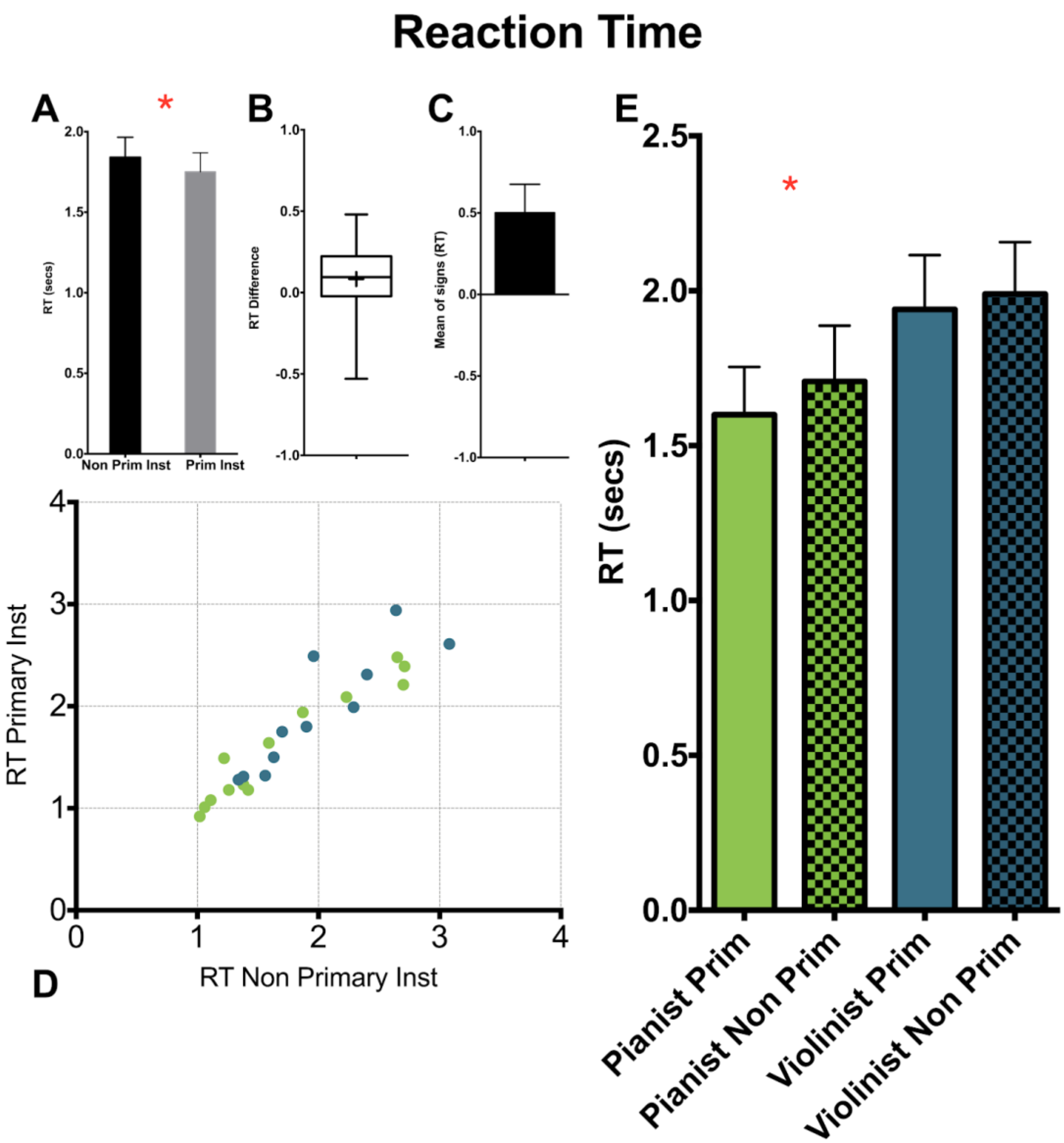

Figure 3.4. The effect of timbral expertise on reaction time. A) Depicts participants' average reaction times for trials in the timbre of their primary and non primary instrument. B) Box plot depicting the average difference in reaction time between trials in the timbre of participants' non primary and primary instrument. C) Depicts the mean of signs (as associated with a sign test) of instances where participants were slower when the trial timbre was in their non primary instrument as opposed to their primary instrument. D) A scatterplot depicting the average reaction time for trials in the timbre of participants' non primary instrument (x-axis) and primary instrument (y-axis). Pianists are represented by green dots and violinists are represented by blue dots. E) Depicts the average reaction time of pianists and violinists for trials in the timbre of their primary and non primary instrument. Error bars represent the standard error of the mean. * $p < 0.05$, ** $p < 0.01$

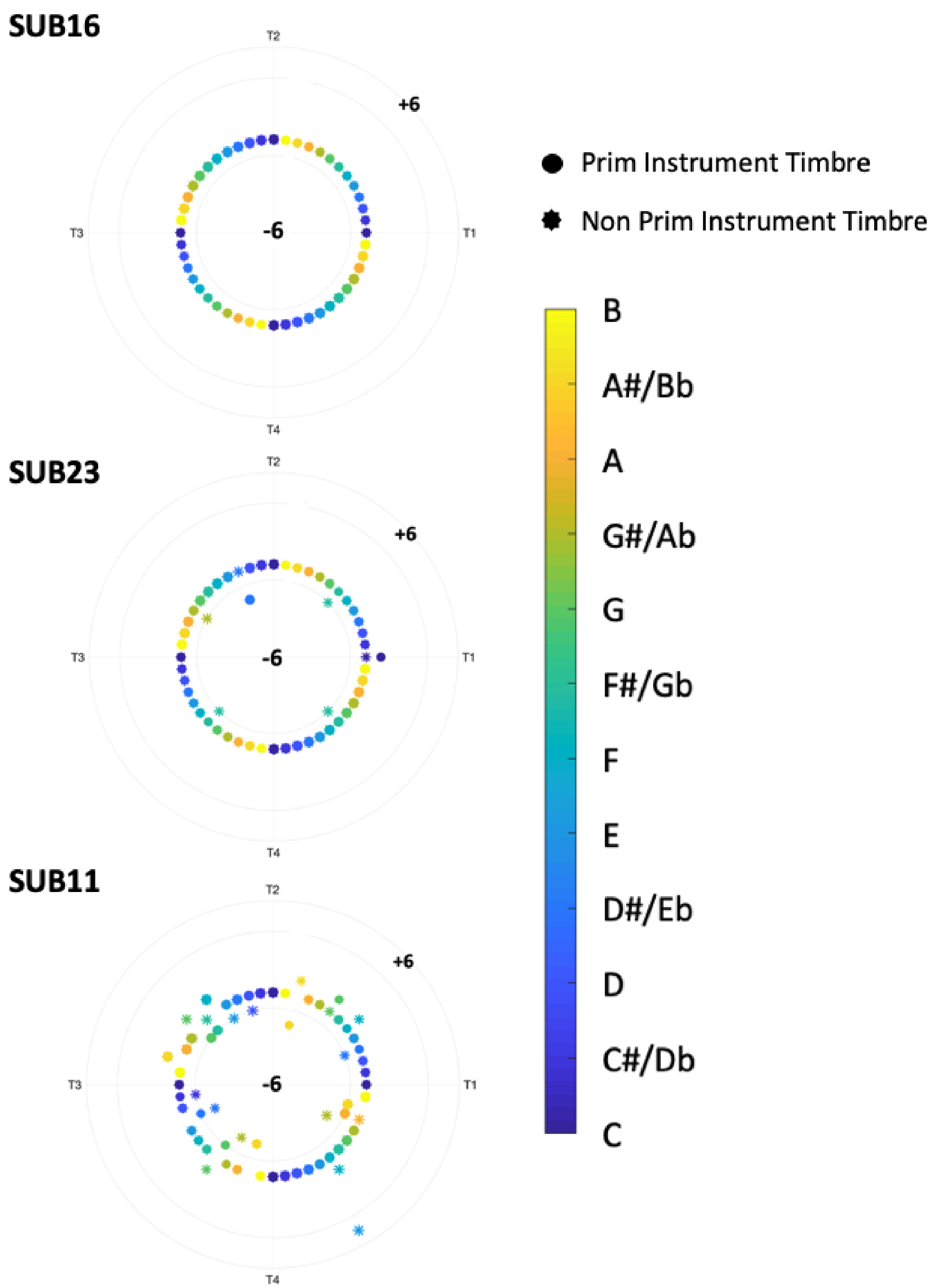

Figure 3.5. Polar plots of errors made during trials in one's primary instrument timbre vs non primary instrument timbre. The data of three participants (SUB16, SUB23 and SUB11) of varying accuracy levels are included. SUB16 did not make any errors during the entire experimental task, while SUB23 made six errors. The ring that each note error falls on corresponds to the number of semitones removed from the correct note (up to 6 in each direction). For each timbre, Quadrants 1 and 2 correspond to the first and second octave of each note, respectively, the first time each note was encountered, Quadrants 3 and 4 correspond to the first and second octave of each note, respectively, the second time each note was encountered.

Chapter 3, is coauthored with Monk, Bradley; Henthorn, Trevor and Deutsch, Diana. The dissertation author was the primary investigator and author of this paper.

GENERAL DISCUSSION

The aforementioned studies on misophonia and AP broadly demonstrate how information associated with certain sounds can impact how we process and respond to them, whether it be emotionally, cognitively or physiologically. While the study in Chapter 1 establishes that there are indeed measurable differences in how those with and without misophonia respond to sounds and other stimuli, the study in Chapter 2 takes things one step further and explores how responses to these sounds can be changed in misophonic individuals, depending on the information paired with the sounds. The study in Chapter 3 examines AP possessors and specifically how information such as timbre can become a part of their mental representation of pitches as a result of experience, to the point where it can facilitate or hinder pitch labeling performance.

The study in Chapter 1 was, notably, the first to apply an experimental paradigm to the study of misophonia. It also utilized semi-structured interviews to help identify the common symptoms, trigger sounds, thoughts, behaviors and coping mechanisms of individuals who suffer from misophonia. The experimental paradigm, which included both self-report ratings and the physiological measure of SCR, was the first to highlight quantitative differences between individuals with and without misophonia in how they subjectively and physically responded to various types of auditory and visual stimuli. The contributions of this study were far-reaching. Prior to 2013, misophonia was largely unknown; many sufferers often felt dismissed or misunderstood by family members and clinicians as most people had never heard of the condition. Publication of this study generated substantial media interest and coverage, which greatly increased the general public's awareness of misophonia, piqued the interest of dozens of other scientists and clinicians, and perhaps most importantly, provided some validation for many individuals suffering from misophonia. As of April 2019, the study has been cited over 90 times, has accrued over 95,000 views and has an attention score that falls in the top 5% of all articles ever scored by Altmetric.

The study in Chapter 2 substantiates a specific finding from the study in Chapter 1 regarding the role of context in the misophonic response while also setting the stage for the development of potential treatments. Through the use of a novel experimental paradigm, we were able to control the amount and type of information that participants received about the sounds they were listening to and as a result, were able to observe how this information influenced how they responded to these sounds. Ultimately, we found that what participants believed about the sounds they were listening to, even if what they believed was incorrect, influenced their response to these sounds. The finding that misophonic individuals were able to find the very same sound to be significantly more or less aversive the next time it was encountered, depending on the information given, was as fascinating as it was encouraging as it empirically demonstrated that there is flexibility in the misophonic response and implies the potential for these responses to be attenuated through reappraisal, perhaps with techniques such as cognitive behavioral therapy.

Some of the limitations of the studies in Chapters 1 and 2 concern the stimuli used and the way they were presented, namely the fact that trigger sounds were generic (and not specific to each participant) and presented on a computer (as opposed to in real life). However, although it may be worthwhile to utilize a more customized set of sounds for each participant in future studies and examine the extent to which reactions to triggers presented on a computer may differ from reactions to triggers presented in real life, it is important to weigh the benefits of both approaches, as the use of consistent sets of computer-presented stimuli minimize the influence of confounding factors and allow us to investigate misophonia in a controlled and methodical manner.

Although the study in Chapter 3 shifts away from misophonia and focuses on AP possessors, it still provides support for the idea that our responses to sounds are influenced by information that has become associated with them. The observed improvement in AP performance shown by AP possessors when labeling pitches in the timbre of their instrument of expertise as opposed to in the timbre of another commonly encountered instrument suggests that AP is more nuanced than assigning a note name to raw frequency information. It also suggests that additional contextual information (such as timbre) and specifically, the extent to which one has associated this contextual information with pitch information (such as through mastering an instrument), can influence AP performance. In other words, AP possessors may utilize a mental template in which they represent pitches in the contexts (timbres, note ranges, etc.) they are most familiar with. Presumably, when assessing a pitch, the more similar it is to the version in their mental template, the easier it will be for them to identify.

This dissertation covered two unrelated auditory phenomena: misophonia, a newly researched condition in which the processing of specific sounds appears to go awry, and absolute pitch, a widely-researched ability in the field of music cognition in which the processing of pitches appears to be enhanced. Through the use of controlled experimental paradigms that highlight the role that contextual information plays in sound processing, the studies in this dissertation contribute to both the early characterization of a relatively unexplored condition and the further characterization of a well-researched ability.