Spring 2016

Phonophobia Mediates the Relationship Between the Myelinated Vagus and Selective Mutism

Alexandra Sherwood Batzdorf
Bard College

Phonophobia Mediates the Relationship Between the Myelinated Vagus and Selective Mutism

Senior Project submitted to

The Division of Science, Mathematics, and Computing

of Bard College

by

Alexandra S. Batzdorf

Annandale-on-Hudson, New York

May 2016

Acknowledgements

This project would not have been possible without the generosity of Dr. Helen Buhler; the instruction of Dr. Lori Perry; feedback from Dr. Amy Winecoff and Dr. Kristin Lane; and, of course, the dedicated guidance, support, and encouragement from my advisor, Dr. Justin Hulbert. Your belief in me was the catalyst that allowed my wild idea to come to fruition.

Thanks are also due to my Los Angeles family members for their invaluable advice and for nurturing my love of learning, as well as to my Bard family for every conversation, bowl of soup, lesson about weather preparation, and instance of trust and acceptance. I could not possibly have been surrounded by better people.

Special thanks go to my starstuff, Camille Thigpen, whose patience, love, and grounding got me through my time at Bard. Tu es éclatant/e.

Words cannot express how grateful I am to each and every one of you for helping me find my voice.

Table of Contents

Abstract

When active, the myelinated vagus (the tenth cranial nerve) acts as a brake that inhibits sympathetic activity by reducing heart rate and blood pressure, and thus allows for social engagement by redirecting metabolic resources. Among those with selective mutism (SM), a disorder characterized by an inability to speak in certain situations, the vagal brake is dysregulated. One consequence of this is a weakening of the middle-ear acoustic reflex (MEAR), which helps clarify human voices and filters out low-frequency background noise, including the speaker's own voice. I tested a proposed etiological model of SM and comorbid social anxiety disorder (SAD) by investigating the relationship between MEAR dysfunction and phonophobia (fear of one's own voice), which were hypothesized to be positively correlated. A nonclinical sample of Bard undergraduate students was recruited. MEARs were assessed using a tympanometer and a signal-to-noise ratio hearing test; phonophobia was gauged by comparing transient anxiety levels before and after reading neutral words aloud. Analyses revealed no reliable correlation between MEAR dysfunction and phonophobia. I suggest an alternative explanation that incorporates a possible compensatory mechanism for unfiltered auditory information. These findings have implications for our understanding of the impact of the vagus nerve on the auditory system, as well as our conceptualization and treatment of SM, which is currently addressed with pharmacological interventions better suited to SAD than SM.

Keywords: selective mutism, vagus, phonophobia, acoustic reflex, social anxiety disorder

Phonophobia Mediates the Relationship Between the Myelinated Vagus and Selective Mutism

Selective mutism (SM) is an anxiety disorder characterized by a consistent inability to speak in certain situations. Such situations are usually unfamiliar or less familiar social situations in which speaking is expected, and are contrasted with other situations in which speaking occurs easily and frequently. The primary recipients of an SM diagnosis are young children; the onset of the disorder is usually before age five. In order for children to receive this diagnosis, their inability to speak must interfere with their lives, last for at least one month (not the first month of school), and not be better explained by discomfort with the language or by a communication disorder. The prevalence of SM is estimated to be between 0.03% and 1%, depending on the population and the ages of its members (American Psychiatric Association, 2013).

Almost all children with SM also receive a diagnosis of an anxiety disorder, usually social anxiety disorder (SAD). SAD is characterized by fear of or anxiety regarding situations with a perceived possibility of public scrutiny. Social situations almost always elicit this response, which can cause individuals with SAD to fear being further scrutinized for reacting with anxiety, thus reinforcing anxious symptoms. In order to receive a diagnosis of SAD, the anxiety must interfere with daily activities, be out of proportion to the situation provoking the anxiety, last at least six months, and not be attributable to other disorders or to medication (American Psychiatric Association, 2013).

Perceived symptomatic overlap of SM and SAD, as well as their extremely high comorbidity—some studies have estimated that as many as 100% of people with SM also have SAD (Muris & Ollendick, 2015)—have led some to view SM as a variant or extreme subtype of SAD, rather than a separate disorder (e.g., Black & Uhde, 1992, 1995; Crumley, 1990; Dummit

et al., 1997). However, not all people with SM are socially anxious. Among those who are, SAD presents even in situations wherein speech is not necessarily expected, and not all distressing social situations result in a failure to speak. Neither SAD nor SM can account for all symptoms of those who have both disorders and the presence of one disorder does not necessitate the presence of the other (Garcia, Freeman, Francis, Miller, & Leonard, 2004). Additionally, there exist developmental and neurological features unique to SM (e.g., Bar-Haim et al., 2004; Heilman et al., 2012; Mash & Wolfe, 2013), and people with comorbid SM and SAD do not demonstrate higher levels of social anxiety than those with just SAD; this discredits the argument that SM is just an extreme subtype of SAD (Henkin & Bar-Haim, 2015; Manassis et al., 2003; Yeganeh, Beidel, Turner, Pina, & Silverman, 2003; Young, Bunnell, & Beidel, 2012). Thus, it does not make sense to claim that SM and SAD are facets of the same disorder.

Despite this evidence, treatment of SM often resembles that of severe SAD. The most commonly prescribed drugs for children with SM are selective serotonin reuptake inhibitors (SSRIs; Carlson, Kratochwill, & Johnston, 1994), which research has demonstrated to be effective treatments for SAD, but there has been very little research regarding their efficacy as treatments for SM. In fact, as of 2008, only 21 studies investigated pharmacological interventions for SM, of which 81% were case studies—the largest sample size was $n = 16$ (all females) and the second largest sample size was $n = 5$ (all males). Forty one per cent of the studies did not examine behavior across multiple settings, and 48% of the studies featured concurrent psychosocial treatments, of which fewer than half varied treatments or included a control group in order to investigate these psychosocial treatments as possible confounds. All but one study used drugs commonly used to treat SAD (most commonly SSRIs, with three instances of monoamine oxidase inhibitors, or MAOIs); the one study that did not use one of these drugs

described an instance in which a child was sedated with nitrous oxide during a dental procedure and was observed to be able to speak while in this sedated state. There were in total 57 participants in all 21 of these studies combined (Carlson, Mitchell, & Segool, 2008). A review of the existing literature since 2008 identified one study that investigated drug treatment of SM: a case study in which a child was administered an SSRI in addition to two other drugs, both of which were prescribed for an unrelated genetic defect (Plener, Gatz, Schuetz, Ludolph, & Kölch, 2012). Clearly, evidence for effective pharmacological interventions for SM is lacking.

In addition to being sparse and poorly operationalized, studies investigating medications that target SM have not produced overwhelmingly conclusive results. In some cases, the drugs being studied did actually reduce SM symptoms, but not significantly more so than placebos (Carlson et al., 2008). Because SM is predominantly diagnosed among children and many cases disappear with age (American Psychiatric Association, 2013), it is possible that the symptom reduction in both placebo and study drug groups was simply a function of time passing and the disorder resolving itself organically. Additionally, rather than using a more objective measure, many studies defined symptom improvement based on the impressions psychiatrists had when meeting children before versus after treatment. Those with SM are much more likely to speak around people they are familiar with (Garcia et al., 2004), so it is entirely possible that any improvement witnessed by psychiatrists can be attributed to increased familiarity with the psychiatrists and thus more comfort speaking around them, rather than to the study drugs. Because of the abundance of methodological flaws in these studies and the lack of statistical power of case studies, it is not possible to draw definitive conclusions regarding the efficacy of commonly prescribed pharmacological interventions for people with SM.

Aside from not being proven to be helpful, these medications can be harmful. For instance, many SSRIs are not recommended for people under the age of 18, since there is an increase in likelihood of suicidal ideation, thoughts of self-harm, and behavioral inhibition as side effects of SSRIs among children; clinicians are therefore advised to deeply consider the costs and benefits of prescribing these drugs to children. Despite this, the only rationale presented for using drugs that help people with SAD as treatments for people with SM is the outdated belief that SM and SAD are symptoms of the same disorder (Carlson et al., 2008). It is therefore necessary to better understand the etiology and neural underpinnings of SM in order to develop informed and effective treatments.

The Polyvagal Theory

One clue to understanding SM involves the polyvagal theory (Stephen W. Porges, 1995), which differentiates between two separate branches (myelinated versus unmyelinated) of the vagus, or tenth cranial nerve. Among vertebrates, the unmyelinated vagus controls shifting between sympathetic and parasympathetic states. The sympathetic state, or "fight or flight," is characterized by a multisystem distribution of resources that allows for quick and efficient escape from dangerous situations by facilitating mobilization; for example, blood rushes to the extremities, heart rate increases, pupils dilate, and bronchi widen. Once an organism is no longer in imminent danger, the parasympathetic nervous system, which is responsible for the state often referred to as "rest or digest," takes over and redistributes resources to processes responsible for homeostatic regulation (Brodal, 2004). When active, the vagus acts as a sympathetic brake, countering the physiological effects associated with fight or flight by slowing the heart rate and dampening the hypothalamic-pituitary axis (HPA). Thus, activation of the vagus, or vagal brake, allows for and promotes parasympathetic activity (Heilman et al., 2012).

Porges (1995) theorized that mammals have a second vagal nerve component, known as the myelinated vagus. The myelinated vagus is crucial to social communication. It originates in the nucleus ambiguous and communicates with special visceral efferent nerves, which are five cranial nerves that innervate striated muscles in the face and head. These are cranial nerves V (which innervates muscles for chewing and middle ear muscles), VII (which innervates middle ear muscles and facial muscles), IX and X (which both innervate laryngeal and pharyngeal muscles), and XI (which innervates neck muscles). Together, these nerves function to control input and output of social information by affecting listening, facial expressions, vocalization, prosody, and head gestures (Stephen W. Porges & Lewis, 2010). Such social engagement is very sensitive to vagal tone. Vagal activity and prosociality demonstrate an inverse U relationship: both abnormally high and abnormally low vagal tone correlate with poor social engagement. This finding suggests that any dysregulation of the myelinated vagus can greatly influence social behavior (Kogan, Oveis, Carr, & Gruber, 2014).

For this reason, a 2012 study investigating vagal function among people with SM may be extremely relevant to the conceptualization of this disorder. The study measured vagal activation and heart rate before, after, and throughout an exercise protocol designed to increase sympathetic activity (i.e., deactivate the vagal brake). Compared to controls, people with SM were found to have a "sluggish" vagal brake, which manifested as less change in vagal activity, less change in heart rate, and a weaker correlation between amount of activity and expected heart rate changes. These results indicate that people with SM demonstrate deficient regulation of the vagal brake and thus difficulty shifting in and out of neurophysiological states. Furthermore, the particularly weak relationship between vagal reactivity and changes in heart rate appears to be unique to those with SM. A dysregulated vagal brake could help explain SM symptoms in that it would

lead to difficulty engaging the larynx and pharynx in order to produce vocalizations, as well as overall difficulty shifting into a parasympathetic, socially oriented state. In addition to enhancing the current understanding of SM, these results suggest the possibility of a biomarker for SM (i.e., vagal dysregulation), which would allow people with or at risk of developing the disorder to be identified, and thus receive treatment, early on in their development (Heilman et al., 2012).

Self-Vocalization and the Stapedius

Another piece of the SM puzzle, which involves abnormal auditory efferent function, is likely also tied to the polyvagal theory. Cranial nerve VII, which is one of the special visceral efferent nerves, innervates a muscle in the middle ear known as the stapedius muscle, which is attached to the stapes bone (Stephen W. Porges et al., 2013). When sound waves reach the human ear, they vibrate against the tympanic membrane, also known as the eardrum. Attached to the eardrum is a chain of three small bones—the malleus, the incus, and the stapes—called ossicles (see Figure 1). The stapes is connected to the cochlea at a site called the oval window. As the eardrum vibrates, it pushes and pulls the ossicles, which in turn push and pull the oval window. Ossicles serve to concentrate and amplify sound vibrations that would otherwise not be forceful enough to displace the fluid within the cochlea. Once vibrations reach the cochlea, the individual frequencies that comprise the vibrations are parsed out and stimulate frequency-specific hair cells, which transduce acoustic signals into electrical impulses. These impulses travel through the eighth cranial nerve to the brainstem, then on to the auditory cortex (Plack, 2014).

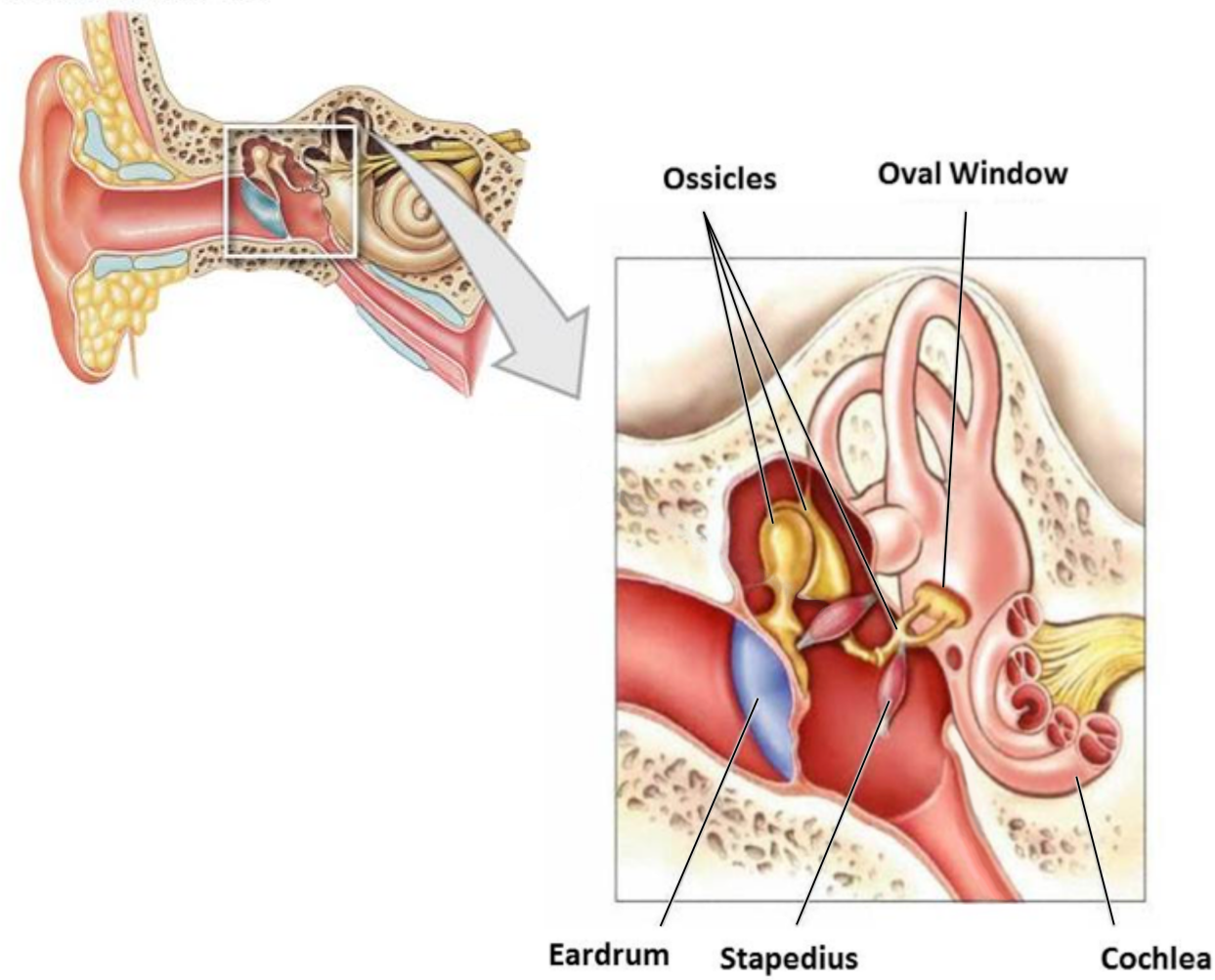


Figure 1. The bones of the middle ear. As sound vibrates against the eardrum, the ossicles (from left to right: malleus, incus, and stapes) transmit this signal to the cochlea. The seventh cranial nerve innervates the stapedius, which is connected to the stapes, the rightmost ossicle. Stapedial contraction stiffens the ossicular chain, thus dampening the transmission of lower frequencies. Adapted from "The Importance of Acoustic Reflex in Speech Discrimination," by K. C. L. de Andrade, S. C. Neto, and P. de L. Menezes, 2011, in I. Ipsic (Ed.), *Speech Technologies*. InTech. Retrieved from http://www.intechopen.com/books/speech-technologies/the-importance-of-acoustic-reflex-in-speech-discrimination and *Fundamentals of Anatomy and Physiology* (10th ed.), by F. H. Martini, J. L. Nath, and E. F. Bartholomew, 2015, San Francisco, CA: Pearson Education, Inc. Copyright 2015 by Pearson Education, Inc.

When loud, sudden, low frequency sounds are transmitted through the ear, the brainstem reflexively sends a signal through the seventh cranial nerve to the stapedius to protect delicate hair cells in the inner ear from damage. Specifically, electrical signals from the cochlea travel through the eighth cranial nerve to the cochlear nucleus, which has projections in both ipsilateral

(i.e., same side of the body as the incoming signals) and contralateral (i.e., other side of the body) superior olivary complexes. The superior olivary complexes stimulate their respective branches of the seventh cranial nerve, which elicit stapedial contraction in both ears (see Figure 2). When the stapedius contracts, it pulls on the stapes so as to restrict its range of motion; this in turn stiffens the entire chain of ossicles, and thus increases the tension of the eardrum. Similarly to how tightening the skin of a drum produces a higher pitched beat, stretching the eardrum creates a filter that emphasizes higher frequency sounds (in the approximate range of 500 Hz to 4000 Hz) and dampens sounds below this range. The process wherein the stapedius muscles contract so as to protect the ears from potentially damaging low frequency sounds is known as the middle ear acoustic reflex, or MEAR (Stephen W. Porges & Lewis, 2010).

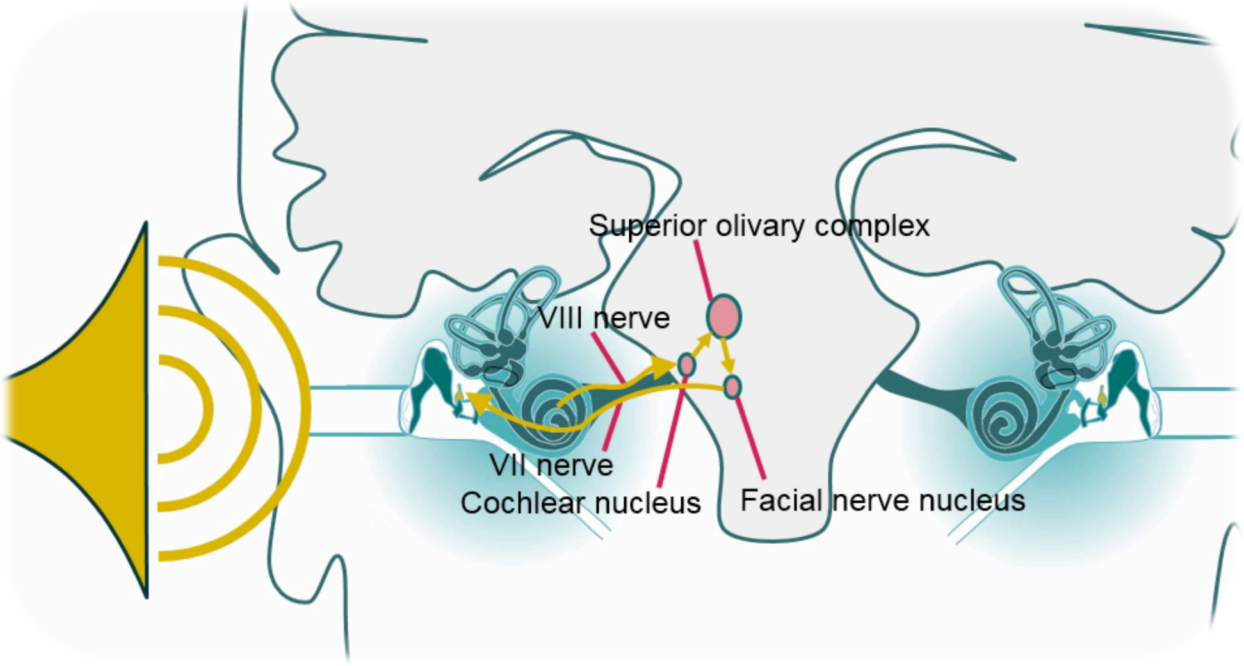


Figure 2. The acoustic reflex arc. Sound vibrations pass through the middle ear and into the cochlea, where they are transduced into electrical signals. These signals travel through the eighth cranial nerve to the cochlear nucleus, then to the superior olivary complexes. The superior olivary complexes stimulate the seventh cranial nerve, which elicits stapedial contraction. Reprinted from *Acoustic reflexes: An introduction*, 2015. Retrieved from https://www.youtube.com/watch?v=3a3Eeuhkh-c

Because the vagus communicates with the seventh cranial nerve (which innervates the stapedius), vagal activity can elicit the MEAR. A likely explanation for this is that sympathetic and parasympathetic states differentially rely on certain frequencies more than others. For instance, low frequency sounds are especially important during sympathetic activation, as anything with a vocalization that is lower than human voices is very likely to be much larger than humans. In a situation that warrants a fight or flight response, such as hearing the low groans of a nearby predator, it would therefore not be beneficial to filter out lower frequencies. When in a social situation, however, the MEAR serves a purpose. Stapedial contraction clarifies subtleties of sounds within the 500 Hz to 4000 Hz frequency band, which makes these sounds become easier to pick out from lower frequency background noise. In humans, as well as all mammals, this emphasized frequency band aligns with that of vocalization. Simply put, stapedial contraction makes it easier for people to make out the sounds of other people's voices (Stephen W. Porges & Lewis, 2010).

In addition to clarifying other people's voices, the MEAR dampens the speaker's own voice. Much of self-vocalization is filtered through bone before reaching the speaker's ears; bone is very absorbent of higher frequencies, which means only the lower frequencies are able to pass through. Without the masking provided by the stapedius, self-vocalizations would sound distractingly loud, strange, and distorted (Henkin & Bar-Haim, 2015). The stapedius is therefore an excellent tool for promoting social engagement. Once there is no more need to exhaust resources in fight or flight, it is no longer necessary to listen for sounds indicating danger. As the vagal brake is initiated, striated muscles integrate to create a cohesive social reaction. For instance, in addition to innervating the stapedius, cranial nerve VII innervates facial muscles. Because of this connection, the movement of looking up to make eye contact simultaneously

activates the stapedius muscle. In addition to disinhibiting the muscles associated with social functioning and allowing for voluntary control, the myelinated vagal brake organically activates muscles implicated in social engagement in conjunction with one another (Stephen W. Porges & Lewis, 2010).

Considering the vagal dysregulation characteristic of those with SM, it should come as no surprise that their stapedius muscles are impaired, as well. Among those with SM, the acoustic reflex threshold (ART), which is the minimum dB level of sound required to elicit the MEAR, is abnormally high or not present at all (Henkin & Bar-Haim, 2015). This is indicative of poor vagal tone and likely plays a prominent role in the development of SM. For instance, sounds in the frequency band emphasized by stapedial contraction often activate the stapedius and are considered to be soothing (Stephen W. Porges & Lewis, 2010). Without effective stapedius muscles, those with SM would be less able to self-soothe. Because stapedial activity increases the perception of subtleties in voices, it is an important tool for recognizing intonation. Decreased recognition of social cues, as well as the accompanying impaired ability to learn and express social signals using intonation, could play a role in the development of SM.

Proposed Etiological Model

In light of these findings, I propose a model (see Figure 3) wherein a sluggish vagal brake—which would result in sustained sympathetic activation, difficulties with social engagement, and hypersensitivity to low frequency noise (Henkin & Bar-Haim, 2015)—would cause people with SM to associate this anxious, predator-wary state and inability to socially engage with their increased attention to their own voices. This would lead them to develop a fear of their own voices, or phonophobia. Phonophobia, along with a general inability to shift into a parasympathetic and socially oriented state, could feasibly lead to the development of social

anxiety. Social anxiety would then be perpetuated as social situations call for speech and social aptitude, which would put more pressure on those with SM, thus stimulating sympathetic activity and strengthening the association between fear and social situations. This is especially likely since social situations would be loud and confusing with an impaired MEAR, which is why people with abnormal MEARs tend to be socially withdrawn and to strongly prefer quiet environments (Bar-Haim et al., 2004).

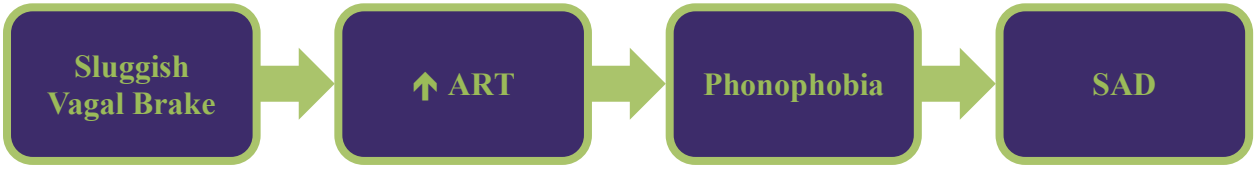


Figure 3. A proposed etiological model of comorbid SAD among people with SM. Because of the dysregulating effect of the vagus on the middle ear acoustic reflex, people with SM would experience simultaneous sympathetic activation and increased awareness of their own voices, which would result in phonophobia. This, coupled with difficulty processing social cues in speech, would lead to the development of SAD. ART = acoustic reflex threshold; SAD = social anxiety disorder.

The main difference between this model and other popular models is the idea that SM could precede SAD, rather than developing as a result of SAD (e.g., Carlson, Mitchell, & Segool, 2008). In order to fill in the gap in the research linking MEAR impairment with phonophobia, I tested the association between ART and anxiety about the sound of one's own voice among members of a nonclinical population. It was hypothesized that there would be a positive correlation between ART and anxiety resulting from hearing the sound of one's own voice.

ART was measured in two ways: directly using a tympanometer and indirectly using a signal-to-noise ratio (S/N) threshold test. Tympanometers are machines that measure impedance, meaning how much of a sound wave is reflected off a surface, rather than allowed through. Since the MEAR pulls the eardrum taut, it greatly increases the eardrum's impedance. This is how it prevents lower frequencies from entering the middle ear. Because of this property, it is possible

to detect whether the MEAR has been triggered based on the impedance of the eardrum. This is accomplished using a probe that rests on the outside of the ear canal and emits four frequency tones (500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz) at varying dB levels, all of which are in compliance with safe noise exposure standards set by the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH; Niquette, 2012). Greater impedance of the eardrum manifests as smaller changes in volume within the ear canal. This is because high compliance (the inverse of impedance) allows for more oscillation of the eardrum, which means greater changes in pressure within the auditory canal as air is pushed and pulled back and forth. The amount of displaced air (which is expressed as volume, usually ml) decreases once the MEAR is elicited. Therefore, ARTs can be determined based on the lowest dB level required to elicit a sharp decrease in the volume of air that is displaced within the auditory canal (Hunter & Shahnaz, 2014). There is essentially no risk involved with the procedure, and since tympanometers are used to detect middle ear pathologies, there is no risk of damaging someone's middle ear as a result of preexisting middle ear pathologies (L. A. Perry, personal communication, November 19, 2015).

Dr. Helen Buhler of Mercy College generously lent a tympanometer for the purposes of this study. Because the tympanometer was available for a limited amount of time, tympanometer data was only collected for a subset of participants. S/N thresholds were therefore examined as the primary measure of MEAR dysfynction. These thresholds measure the ability to understand speech (the signal) within the context of varying levels of background noise (the noise). Because the MEAR clarifies speech and dampens background noise (Stephen W. Porges & Lewis, 2010), S/N threshold was expected to be an indirect measure of ART. Furthermore, S/N threshold has

been found to be positively correlated with ART (e.g., Harkrider & Smith, 2005; Henkin & Bar-Haim, 2015; Porges et al., 2013).

In order to measure anxiety resulting from the sound of one's own voice, participants were instructed to fill out a questionnaire rating state-dependent (i.e., transient) anxiety levels before and after reading a list of neutral words aloud. The changes in their responses as a result of hearing their own voices served as the measure of phonophobia. Because a high ART, which would indicate low stapedial sensitivity, would make one's own voice seem distractingly loud, and because a high ART likely indicates concomitant difficulty shifting out of a sympathetic state, it was hypothesized that S/N threshold would be positively correlated with anxiety resulting from hearing one's own voice. Secondary hypotheses were that ART (measured via tympanometer) would also be positively correlated with anxiety resulting from hearing one's own voice for the same reason, and that these two correlations would not be significantly different, as ART and S/N threshold were both expected to reflect MEAR function.

This study clarifies the relationship between ART and phonophobia, thus elucidating the feasibility of the proposed model. In turn, the developmental course of SAD among those with SM can be better understood. If results are in line with the hypotheses, this would provide evidence in support of the proposed model; if results are not in line with the hypotheses, this could suggest that an alternative explanation should be explored, such as a possible mechanism to compensate for irrelevant auditory stimuli that are not adequately filtered by the middle ear. Findings will improve our understanding of SM, which will allow for the possibility of more appropriate, research backed treatments for SM.

Method

Participants

Forty-five undergraduate students at Bard College participated in the study. Participants were recruited via flyers hung around the Bard College campus (see Appendix A); messages posted on Facebook and included in Bard Daily Mail email blasts; short, verbal descriptions of the study announced at the beginning of classes; and an information table in the Bard Campus Center, where I sat with a banner advertising the study and asked passersby if they would like to participate. No demographic information about participants was collected beyond asking each participant if they met inclusion criteria, as neither gender nor age affects ARTs (Gates, Cooper, Kannel, & Miller, 1990; Hunter & Shahnaz, 2014; Ludewig et al., 2003; Rawool, 1998) or S/N thresholds (Kallen et al., 2012). All participants were given a piece of candy and an entry into a raffle to win a $50 Amazon gift card, which was awarded to one randomly chosen participant.

In order to be eligible for the study, participants had to be fluent in English (to ensure comprehension difficulties did not impact their comfort reading words aloud or their ability to understand and repeat S/N threshold test items); be at least 18 years old; and have normal hearing. Although it is possible that some participants unknowingly had mild hearing loss, this would not affect ARTs (L. A. Perry, personal communication, November 19, 2015) or S/N thresholds (Ferman, Verschuure, & Van Zanten, 1993; Kallen et al., 2012; Zecker et al., 2013).

Of the 45 people who participated in the study, two were excluded—one because they did not follow STAI directions properly (after completing one of the STAI versions, they reported that they had mixed up whether the numeric anchors indicating degree of anxiety were ascending or descending partway through filling out that STAI version), and one because they did not believe that their language comprehension was adequate for the S/N threshold task (after

completing the task, they reported feeling that their lack of familiarity with English made adjusting to the task especially difficult; indeed, their scores improved as the task increased in difficulty). Of the remaining 43 participants, ARTs alone was collected from 10 people, S/N thresholds alone were collected from 12 people, and both ARTs and S/N thresholds were collected from 21 people (total ARTs: $n = 31$; total S/N thresholds: $n = 33$). The tympanometer was unable to detect one person's left ear ART and one person's right ear ART. Additionally, two participants did not demonstrate detectable right ear MEARs at the highest dB level presented. Their right ear ART data was therefore not included in analyses. All told, analyses included 33 participants' S/N thresholds, 30 participants' left ear ARTs, and 28 participants' right ear ARTs.

Materials

Modified State-Trait Anxiety Inventory (STAI) Form Y-1. The state section of the STAI Form Y (STAI-Y-1) is a sensitive, valid, reliable, highly utilized inventory for detecting shifts in transient anxiety levels (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) that has in the past been used to measure anxiety specifically related to speech (Jeger & Goldfried, 1976; Lamb, 1972; Lent, Russell, & Zamostny, 1981; Slutsky & Allen, 1978). STAI Form Y is an updated version of Form X that is preferred to its predecessor due to Form Y's clearer, less outdated wording and improved psychometric properties (Spielberger & Sydeman, 1994). In the present study, STAI-Y-1 was administered twice in order to capture changes in anxiety as a result of hearing one's own voice. So as to prevent a carryover effect, this 20-item questionnaire was divided into two 10-question forms. All items have demonstrated high internal consistency, and psychometric properties remained when the scale was divided up in the past (Bayrampour, McDonald, Fung, & Tough, 2014; Chlan, Savik, & Weinert, 2003; Court, Greenland, &

Margrain, 2010; Tluczek, Henriques, & Brown, 2009). For the purposes of this study, the form that contains the first 10 questions of STAI-Y-1 is referred to as "STAI Version A," and the form that contains the last 10 questions of STAI-Y-1 is referred to as "STAI Version B."

Participants rated each item using a Likert scale ranging from 1 (*not at all*) to 4 (*very much so*). Within each STAI version, five items—including "I am tense" and "I am presently worrying over possible misfortunes"—indicated the presence of anxiety, while five reverse coded items—including "I am relaxed" and "I feel steady"—indicated the absence of anxiety (see Appendix B). As recommended in the STAI scoring manual (Spielberger et al., 1983), the numeric anchor circled for each item (or the reverse code of the circled number, when applicable) was added up to determine both the initial STAI score and the post-reading activity STAI score. Change in anxiety (Δ STAI) scores were calculated by subtracting participants' initial STAI scores from their post-reading activity STAI scores, as is most commonly done to determine Δ STAI scores (e.g., Consonni et al., 2010). Within each STAI version (i.e., STAI Version A and STAI Version B), the highest possible score (indicating most anxiety) is 40, and the lowest possible score (indicating least anxiety) is 10. Therefore, Δ STAI scores range from ±30 (indicating greatest change in anxiety) to zero (indicating no change in anxiety); a positive Δ STAI score indicates an increase in anxiety after the reading activity, while a negative Δ STAI score indicates a decrease in anxiety after the reading activity.

Neutral words list. Words with high positive or negative valence have been shown to influence attention, memory, mood, prefrontal cortical and amygdala activity, and linguistic processing; these effects can interfere with tasks and can prime people to gravitate toward words with similar valences later on (Soares, Comesaña, Pinheiro, Simões, & Frade, 2012; Stadthagen-Gonzalez, Imbault, Pérez Sánchez, & Brysbaert, 2016). In order to ensure that participants

completed the post-reading activity STAI based on their reactions to hearing their own voices, rather than their reactions to encountering particularly positive or negative words, they were presented with only neutral words during the reading activity. Neutral words were selected from the Affective Norms of English Words (ANEW) database (Bradley & Lang, 1999). ANEW is a widely used, highly reliable catalogue of American English words' affective characteristics, which have been demonstrated to correspond tightly to affect-related physiological responses, such as skin conductance and heart rate (Soares et al., 2012).

ANEW reports each word's mean valence as rated on a 9-point Likert scale (the Self-Assessment Manikin), with 1 being *most unpleasant* and 9 being *most pleasant* (Bradley & Lang, 1999). As recommended by Ferré, Guasch, Moldovan, and Sánchez-Casas (2012); Warriner, Kuperman, and Brysbaert (2013); and Ho et al. (2015), a word was considered to have neutral valence if its mean was between 4 and 6. All such words (231 total) were compiled. Of these 231 neutral valence words, 100 were removed based on researcher's discretion because they did not seem neutral. For some of these removed words, it was suspected that multiple definitions led some people to rate them as positive and some to rate them as negative, creating a wide distribution of valence ratings with a misleadingly neutral mean[1]. Other removed words included those that were expected to be especially likely to vary in valence from person to person (e.g., body parts, animals, common allergens, foods, recreational drugs, religious symbols, sexual acts, weapons). After these 100 words were removed from the compiled list of neutral words, 131 words ($M = 5.28$, $SD = 0.39$) remained. These 131 words (see Appendix B) were randomly ordered using a list randomizing website (Haahr, n.d.) so that participants would not be

[1] For example, the word "fall" could be interpreted as the season of autumn, which might have positive connotations for many raters, or as the result of tripping, which might have negative connotations for raters.

distracted by irrelevant list characteristics, such as an alphabetical pattern.

No study was found in existing literature that details how experimenters chose how many words should be read aloud. However, in a study that used real time feedback to create the illusion among participants that foreign-sounding voices were their own voices, trends indicate that participants would adopt the voice as their own by about 127 words (Zheng, MacDonald, Munhall, & Johnsrude, 2011). Another study, which examined self-ratings of verbal performance among people with social anxiety, had participants read 107-word stories aloud (Hirsch & Clark, 2007). Holzman and Rousey (1966) looked specifically at how participants reacted to hearing their own voices played back on a tape recorder. They had participants speak for about a minute, but found that people reacted within about five seconds of hearing their own voices played back. Likewise, a study that examined functional brain differences between hearing one's own voice, hearing a distorted version of one's own voice, and silently reading along while someone else reads aloud used sets of words that were about a minute long (McGuire, Silbersweig, & Frith, 1996). A study that sought to investigate participants' subjective and physiological reactions to their own tape-recorded voices instructed participants to speak for seven seconds, which was enough to elicit both subjective and physiological responses—including changes in pulse amplitude, which is a measure of vagal activity (S W Porges, Doussard-Roosevelt, Portales, & Greenspan, 1996). However, the researchers recommend waiting about 30 seconds before counting results, so that the novelty of the task wears off. They also report a pattern wherein participants react defensively to hearing their own voices, then attempt to rationalize this reaction, and finally try to pinpoint the specific aspects of their voices that prompted this reaction; this entire process takes about two minutes (Holzman, Rousey, & Snyder, 1966). The average number of words spoken per minute is typically between 120 words and 180 words

(Miller, Maruyama, Beaber, & Valone, 1976). However, because reading a list of words aloud is more cumbersome and unnatural than conversational speech, participants in the present study took about 2 minutes and 13 seconds to read the 131 neutral words aloud. Therefore, while in the approximate range of the number of words in most studies involving reactions one's own voice and reading aloud, 131 words is on the higher end. This conservative selection was made to ensure that all participants had ample time to react to hearing their own voices.

dB Volume iPhone application. To rule out the possibility that some participants might compensate for anxiety resulting from hearing their own voices by speaking at a lower volume, the iPhone decibel meter application dB Volume (2016) was used to monitor sound levels while participants read the neutral words list out loud. dB Volume has been demonstrated to detect sound levels within the range of human speech accurately and reliably (Keene et al., 2013; Nast, Speer, & Le Prell, 2014). The iPhone microphone was placed approximately 3 ft away from participants, as is recommended for decibel meters (Keene et al., 2013; Nast et al., 2014).

Self-Assessment Manikin (SAM). In order to ensure that participants actually perceived the neutral words to be neutral, after filling out the post-reading activity STAI, participants were asked to rate the pleasantness of the overall content of the words they read aloud. Twenty-six people did so using the SAM valence scale (see Appendix B), a 9-point Likert scale ranging from 1 (*very unpleasant*) to 9 (*very pleasant*) with an accompanying pictorial representation at five of the nine points (i.e., at every other point; Bradley & Lang, 1994). SAM has demonstrated excellent reliability and validity (Backs, da Silva, & Han, 2005; Bradley & Lang, 1994; Morris, 1995) and was chosen in part because it was also used to norm the words in the ANEW database from which the neutral words list was drawn (Bradley & Lang, 1999).

While awaiting approval from the Bard Institutional Review Board for an amendment

that included using SAM, participants (n = 17) were administered a Likert scale with the same labeled anchors as SAM, but without pictures (see Appendix B). Participants were instructed to circle either one of five numbered anchors, in line with Jalenques, Enjolras, and Izaute (2013), or one of four dots that served as anchor points between the five numbers. This valence scale was therefore a 9-point scale, much like SAM. Although it did not feature drawings, researchers have found that scales with verbal anchors alone yield results equivalent to those obtained using both verbal and pictorial anchors (Bradley & Lang, 1994; Stadthagen-Gonzalez et al., 2016; Warriner et al., 2013). Furthermore, a two-tailed Welch's approximate t-test determined that the difference between valence ratings among those who completed this scale (M = 5.76, SD = 1.20) and those who completed SAM (M = 5.46, SD = 1.07) was not statistically significant; $t(31.43) = 0.85$, p = .40. It is therefore unlikely that valence ratings were affected by which of the two valence scales each participant received.

National Institutes of Health Toolbox Words-in-Noise Test (WIN). WIN, a measure of S/N thresholds, was administered via the National Institutes of Health Toolbox iPad application (NIH TB app; 2016). Participants listened to a total of 70 target words (35 per ear), none of which overlapped with those on the neutral words list, at decreasing ratios of target word- (signal) to-background babble (noise). There were seven such target word-to-background babble ratios: 24 dB, 20 dB, 16 dB, 12 dB, 8 dB, 4 dB, and finally 0 dB S/N. The first ratio (24 dB S/N) was the easiest to understand and the last ratio (0 dB S/N) was the hardest to understand. Each ratio level featured five target words. WIN used one of two 35-word lists for each ear. Which list corresponded to which ear, as well as which ear was tested first, was automatically randomized by the NIH TB app. Within each of the two lists, the five words that corresponded to each S/N ratio level remained constant; however, the order of these five words

within each S/N ratio level was randomized. For example, 20 dB S/N of List 2 might present the words “haze, such, tire, shawl, gun” for one participant, then “gun, such, shawl, haze, tire” for the next; the 20 dB S/N ratio level of List 2 always had those five words, but the order changed (Wilson & Burks, 2005). Although the order of the target words within each level changed, the segment of background babble that accompanied each word stayed constant, so as to prevent noise caused by an interaction between target words and corresponding background babble (Wilson, 2003).

Participants listened to WIN through Sennheiser HD201 headphones, as recommended by the NIH TB app (“NIH Toolbox Administrator’s Manual,” 2012), and repeated each target word out loud. Their S/N thresholds were defined as the S/N ratio level at which they correctly repeated 50% of the target words. The NIH TB app automatically generates the S/N threshold for each ear based on how many target words participants correctly identified. S/N thresholds can range from 26 dB S/N, indicating worst possible performance, to -2 dB S/N, indicating best possible performance. WIN has demonstrated excellent reliability and validity (Gershon et al., 2013; Wilson, Carnell, & Cleghorn, 2007; Wilson & McArdle, 2007; Zecker et al., 2013), is not affected by ambient noise, and was developed specifically to be accessible to people without audiological training (Zecker, 2012; Zecker et al., 2013). I was trained to administer WIN by an NIH TB app e-learning module (National Institutes of Health & Northwestern University, 2015) in conjunction with the NIH TB Administrator’s Manual (2012).

Maico Otowave 102 Hand Held Portable Tympanometer. This tympanometer, generously loaned by Dr. Helen Buhler of Mercy College, has a current calibration certificate and was used to obtain all ART data. The Maico Otowave 102 Tympanometer is the tympanometer model recommended by the National Institutes of Health (Zecker et al., 2013).

Ipsilateral ART measurements were obtained at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz, but only 2000 Hz thresholds were included in data analysis, as this frequency produces the broadest range of measurements (see: Bar-Haim et al., 2004; Muchnik et al., 2013). Each frequency was initially presented at 85 dB and increased in 5 dB increments until a threshold was detected. In compliance with safety standards, the maximum sound pressure level was 100 dB for 500 Hz, 1000 Hz, and 2000 Hz frequencies and 95 dB for the 4000 kHz frequency (*Maico Otowave 102 Hand Held Portable Tympanometer operating manual*, n.d.; Niquette, 2012). ART measurements were performed three times per ear at each frequency, and ART was defined as the lowest sound pressure level required to elicit a change in compliance of at least 0.03 ml in at least two of the three trials, as is standard (e.g., Bar-Haim et al., 2004; da Cruz Fernandes, Momensohn-Santos, Martins Carvalho, & de Queiroz Carvalho, 2013; Day & Feeney, 2008; Muchnik et al., 2013). Ear tips were thoroughly cleaned with alcohol wipes after each use. I was trained to perform ART testing by Dr. Lori Perry, Au.D. and via an online audiology training module (Mack, 2011).

Procedure

All participants were run in Preston 121. Participants scheduled appointments online to take part in the study (http://bit.do/PsychStudy). They were greeted in the lobby of Preston and escorted to the testing room, where they were given information about the study and their rights as research participants in verbal and written form (see Appendix C). After giving informed consent, participants were instructed to fill out one of the STAI versions based on how they felt in that moment. Participants were assigned ID numbers chronologically based on when they participated; those with odd ID numbers (e.g., 001, 003, 005, etc.) completed STAI Version A before the reading activity and STAI Version B after the reading activity, and those with even ID

numbers (e.g., 002, 004, 006, etc.) completed STAI Version B before the reading activity and STAI Version A after the reading activity. This was to ensure Δ STAI was not determined by the order in which each STAI version was filled out. Each STAI version took about two minutes to complete.

After filling out the first STAI version, participants were asked to stand with each foot on either side of a piece of orange tape on the floor. This was to make sure they were 3 ft from the dB Volume iPhone application. Participants were told they would be given a list of words to read aloud, and that they should read them from left to right at a volume as if speaking to someone on the other side of the room. They were told to speak at a normal pace, and were given the example of "word… word… word… word…" at a rate of about one word per second. Participants were then told not to worry about pronunciation, and that I would be monitoring an application that measures the acoustic properties of the room (dB Volume iPhone application), but does not record their words. The comments about pronunciation and close application monitoring served to help participants avoid feeling as though their performance was being evaluated, so that their Δ STAI scores would not capture stage fright or performance anxiety. Before receiving the list of neutral words, participants were told that another questionnaire (post-reading activity STAI) would be waiting for them after the reading activity, and that the same instructions would apply—they should fill it out based on how they feel at that moment. Participants received these instructions before the reading activity so that they would fill out the second STAI directly after hearing their own voice.

After filling out the post-reading activity STAI, participants filled out SAM (or the equivalent scale) based on the content of the words they read aloud. When applicable, participants then moved on to WIN. As recommended by WIN (Wilson, 2003), participants were

told that they would hear a voice instructing them to repeat various words, and that they should repeat these words out loud so that I could hear them; they should be sure to enunciate. They were then given an example: if the voice says, “Say the word cat,” they should say “cat.” Participants were then told that if they were unsure what the voice said, they should say what they think they heard, or if they had no idea what the voice said, they should say nothing—the app would move on to the next word automatically, rather than waiting for them to say something. Finally, participants were told that they would hear people talking in the background, but they should do their best to ignore them and focus on the main voice asking them to repeat the words, and that they would hear words in only one ear, and then in only the other ear. At this point, participants were asked if they had any questions, then if they were ready to begin.

For each S/N ratio level, the iPad screen—which was visible to me, but not to participants—displayed the five words that participants would hear. As participants repeated each word, I clicked on the corresponding word if it had been repeated correctly. Because feedback about one’s performance can influence one’s future performance (e.g., Fishbach, Eyal, & Finkelstein, 2010; Fishbach & Finkelstein, 2011; Osinsky, Seeger, Mussel, & Hewig, 2016; van de Vijver, Ridderinkhof, & de Wit, 2015), I clicked an inactive part of the iPad screen when participants repeated a word incorrectly, so that they could not gauge their performance. For the same reason, I maintained a neutral facial expression and kept my eyes on the screen throughout the measure. WIN took about six minutes to administer.

When applicable, participants then went on to have their ARTs measured. They were instructed to sit still and not speak, unless they wanted me to stop at any time. Then, they were told that they would hear strange beeps in each ear three times. This process took about three minutes. After their ARTs were determined, they received a verbal and written debriefing

statement (see Appendix C). Those whose data were collected before WIN was approved by the Institutional Review Board were given their ID numbers and were told to bring them if they choose to come back to complete WIN, as no identifying information was associated with their ID numbers in any way. They were then thanked for their time and told to select a piece of candy.

Data Analysis

All statistical tests were two-tailed and the critical value was set at $\alpha = .05$, as is most common (Aron, Coups, & Aron, 2013, p. 113; Whitlock & Schluter, 2008, p. 136). Data were analyzed using R Version 3.2.3. Comparisons were made between sample scores and population norms using Welch's approximate *t*-test, as equal variance could not be assumed. Left ear and right ear data were compared using the paired *t*-test, and their correlations were compared using Meng's *Z*-test, as is appropriate for comparing overlapping correlations based on dependent groups (Meng, Rosenthal, & Rubin, 1992). WIN scores and Δ STAI scores, ARTs and Δ STAI scores, and WIN scores and ARTs were compared using tie-corrected Spearman's r_s correlations. This is because ARTs did not appear to be normally distributed, and because both Likert scales (i.e., Δ STAI) and the dB scale (which is logarithmic) are ordinal, which Spearman's r_s is better equipped to handle (see: Blumberg, Cooper, & Schindler, 2011; Mukaka, 2012). Because of the discrete nature of the ART measurement, a median split was performed (e.g., Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015) to compare ARTs among those with high Δ STAI scores with ARTs among those with low Δ STAI using a Welch's *t*-test (again, equal variance could not be assumed); WIN scores were compared this way as well for continuity. The median split of Δ STAI scores was also used to evaluate dB Volume levels with a Welch's *t*-test.

Finally, the correlation between ART and Δ STAI was compared with the correlation between WIN and Δ STAI with Meng's *Z*.

Results

Unless otherwise stated, there were no significant differences between left ear data and right ear data in any analysis. Therefore, for the sake of concision, the data from only one ear is reported, as is commonly done (e.g., Laukli & Mair, 1980; Nondahl, Cruickshanks, Wiley, Tweed, & Dalton, 2013). Left ear data were arbitrarily chosen for analysis; see Appendix D for right ear data.

Data from the present study were compared with normative data for the STAI (Spielberger et al., 1983), the valence of the neutral words list (Bradley & Lang, 1999), WIN scores (Kallen et al., 2012), and ARTs (da Cruz Fernandes et al., 2013). Welch's approximate *t*-test assumes that samples come from normally distributed populations (Whitlock & Schluter, 2008, pp. 292–293). The sample data for all measures except ART appear to come from normally distributed populations (see histograms below); ART data represent only four discrete dB values, so it cannot necessarily be assumed that they come from a normally distributed population. However, Welch's approximate *t*-test is very robust, as even if the population distribution is not normal, the central limit theorem states that sample distributions will approach normality as they increase (Whitlock & Schluter, 2008, pp. 245, 325). There were no significant differences between data from this study and normative data on any measure (see Table 1), meaning the scores of the participants that comprise this study's sample are in line with those reported by prior literature.

Table 1

Means from the Present Study Compared with Normative Data

	Sample Data		Normative Data			
Measure	*n*	*M* (*SD*)	*n*	*M* (*SD*)	*t* (*df*)	*p*
STAI	43	36.77 (9.10)[a]	777	37.89 (11.3)	-0.78 (49.46)	.44
Word valence	43	5.58 (1.12)	131	5.28 (0.39)	1.75 (45.44)	.09
WIN score	33	4.88 (1.48)[b]	53	4.79 (4.07)	-0.15 (71.26)	.88
ART	30	90.33 (4.90)[b]	20	90.00 (5.34)	-0.22 (38.39)	.82

Note. Data from the present study (sample data) did not differ significantly from normative data on any measure. Word valence = SAM ratings of neutral words list.
[a]Mean and *SD* were determined by doubling initial STAI scores, as normative data reflect baseline anxiety levels. [b]Sample data are reported for left ear only. However, WIN scores as determined by each participant's better ear score ($M = 4.11$, $SD = 1.57$), as reported in normative data[2], and right ear ARTs ($M = 90$, $SD = 4.91$) are also not significantly different from normative data; $t(73.10) = -1.09$, $p = .28$ and $t(38.42) = 0$, $p = 1.0$, respectively.

Correlation Between WIN Scores and Δ STAI Scores

In order to investigate the relationship between the MEAR and anxiety resulting from hearing one's own voice, I ran a Spearman's r_s correlation between WIN scores and Δ STAI scores. WIN score data ($n = 33$, $M = 4.88$, $SD = 1.48$; see Figure 4a) and Δ STAI score data ($n = 43$, $M = 1.28$, $SD = 4.26$; see Figure 4b) do not violate assumptions of Spearman's r_s, as Spearman's r_s is a nonparametric test and therefore does not make assumptions about a population's distribution (Whitlock & Schluter, 2008, pp. 342, 443). There was no significant correlation between WIN score and Δ STAI score, $r_s(31) = -.05$, $p = .79$ (see Figure 5), meaning the hypothesis that higher ARTs (indicated by higher WIN scores) would be positively correlated with greater anxiety resulting from hearing one's own voice (indicated by higher Δ STAI scores) was not supported.

[2] Because participants have a WIN score for each ear, Kallen et al., 2012 chose to report each participant's better ear score (i.e., whichever ear had a lower WIN score).

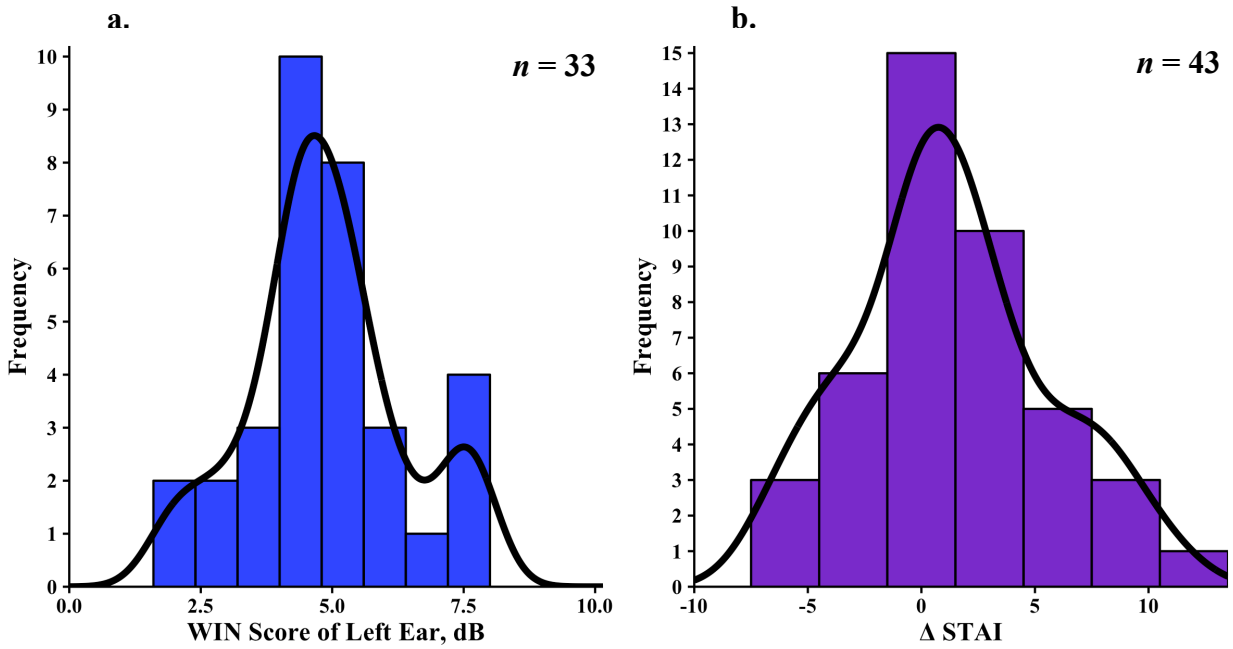


Figure 4. Distribution of left ear WIN scores (a) and Δ STAI scores (b). These data appear to come from a normally distributed population. Frequency = number of participants with each score.

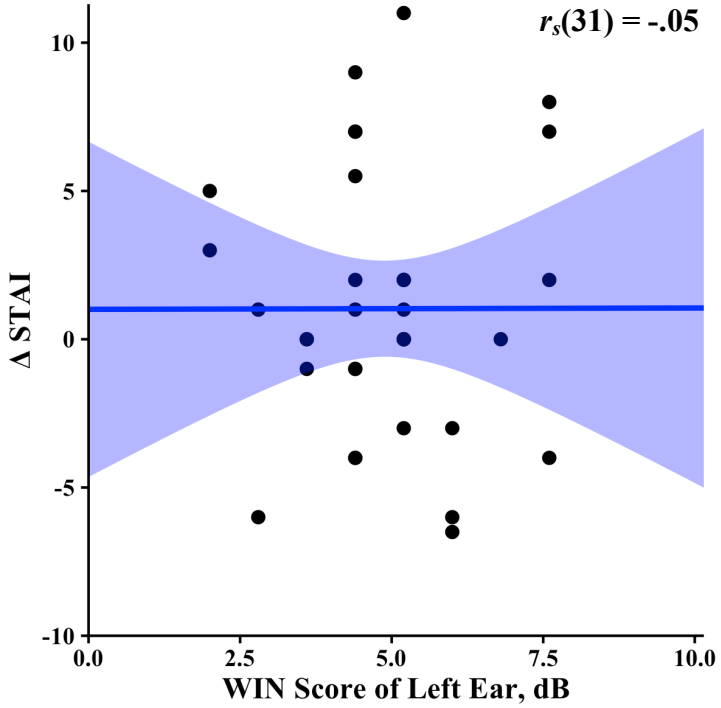


Figure 5. Relationship between left ear WIN score and Δ STAI score. No correlation was found. Error cloud represents 95% confidence interval.

Exploratory Analyses

Tympanometer data was only collected from a subset of participants due to restricted access to the device; still, the data it yielded (ARTs) and their correlation with Δ STAI scores could provide a preliminary sense of the relationship between the MEAR and anxiety resulting from hearing one's own voice. Additionally, ART data can be used to evaluate the validity of WIN as an indirect measure of the MEAR.

Correlation between ART and Δ STAI score. Because there were only four points of measurement (85 dB, 90 dB, 95 dB, and 100 dB), it is difficult to determine the exact nature of the population distribution of ARTs. The sample distribution (n = 30, M = 90.33, SD = 4.90; see Figure 6) does not appear normal, but seems truncated. In support of this claim, other studies with wider ranges of measurements report normal distributions of ARTs (e.g., Laukli & Mair, 1980). It is therefore possible that the population distribution of ARTs is normal. There was no significant correlation found between left ear ART and Δ STAI score; $r_s(28) = .11$; $p = .55$ (see Figure 7). Left ear ART data therefore do not support the hypothesis that ART and Δ STAI would be positively correlated[3].

[3] A positive correlation was found between right ear ARTs and Δ STAI scores ($r_s(26) = .38$, $p = .04$); however, one should take caution in interpreting these results (see Discussion).

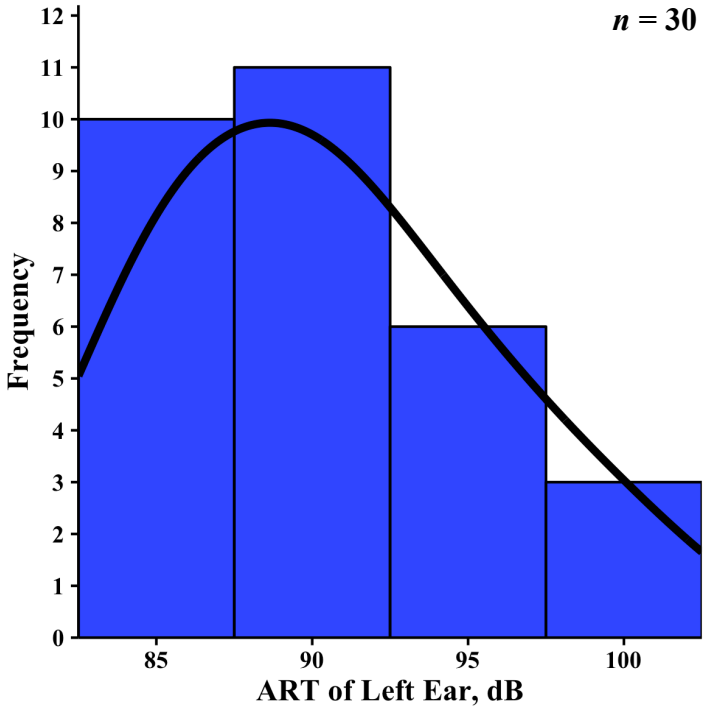


Figure 6. Distribution of left ear ARTs. This distribution is truncated, but could come from a normal population. Frequency = number of participants with each dB level ART.

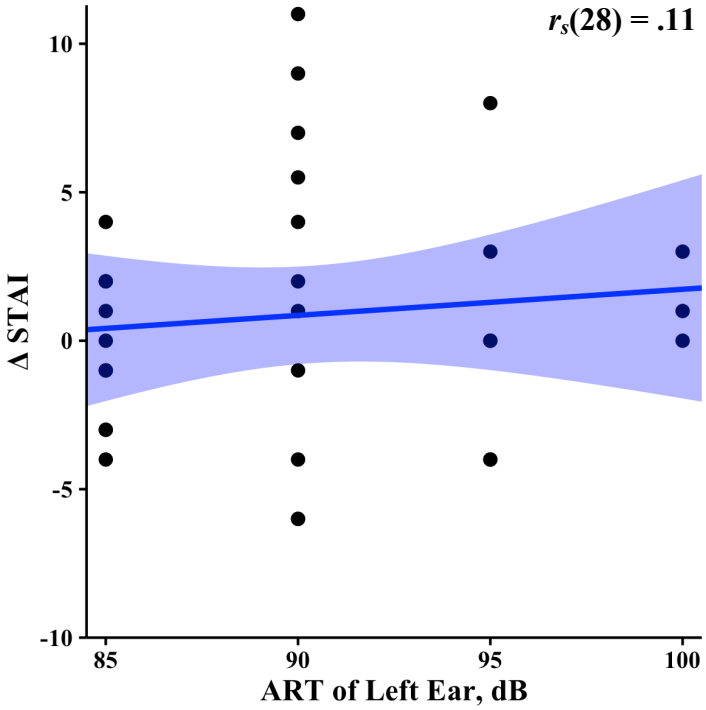


Figure 7. Relationship between left ear ART and Δ STAI score. No correlation was found. Error cloud represents 95% confidence interval.

Because only four points of measurement were possible using the tympanometer, a median split of Δ STAI scores (e.g., Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015) was conducted to attempt to clarify the relationship between ARTs and Δ STAI scores. A Welch's *t*-test comparing ART scores in the high Δ STAI scores group (i.e., Δ STAI score greater than median; $n = 17$, $M = 90.29$, $SD = 5.44$) and ART scores in the low Δ STAI scores group (i.e., Δ STAI score less than median; $n = 13$, $M = 90.38$, $SD = 4.31$) were not significantly different, $t(27.94) = 0.05$, $p = .96$ (see Appendix E, Figure E1). This corroborates the finding that there was no correlation between Δ STAI scores and left ear ARTs, and thus does not support the hypothesis that higher ARTs would be positively correlated with greater anxiety resulting from hearing one's own voice.

For continuity, a median split of Δ STAI scores was used to compare WIN scores, as well. Left ear WIN scores in the high Δ STAI scores group ($n = 18$, $M = 4.98$, $SD = 1.22$) and in the low Δ STAI scores group ($n = 15$, $M = 4.77$, $SD = 1.79$) were not significantly different, $t(24.04) = 0.38$, $p = .71$. This corroborates the above findings; this alternative way of looking at the data did not support the hypothesis that higher WIN scores would be positively correlated with greater anxiety resulting from hearing one's own voice.

This median split was also used to analyze dB Volume data. To rule out the possibility that participants were compensating for a fear of their own voices by speaking quietly, a Welch's approximate *t*-test was conducted to compare dB Volume levels within the low Δ STAI scores group and within the high Δ STAI group. There were no differences in dB Volume scores (see Appendix F), meaning participants were likely not compensating for a fear their own voices by speaking at a softer level.

Comparison of ART and WIN score. Meng's *Z* was used to compare the correlation between WIN scores and Δ STAI scores with the correlation between ARTs and Δ STAI scores. WIN scores and ARTs themselves were not correlated ($r_s(18) = .19$, $p = .42$). Their respective correlations with Δ STAI scores were also not related, $z = -0.28$, $p = .78$. This does not support the hypothesis that WIN scores and ARTs would both be valid constructs for detecting the MEAR; this finding contradicts those of other studies, which have reported finding a positive correlation between S/N threshold scores and MEAR dysfunction (e.g., Harkrider & Smith, 2005; Henkin & Bar-Haim, 2015; Porges et al., 2013).

Discussion

The present study sought to investigate the relationship between the MEAR and phonophobia, so as to bridge the gap in a model that proposes an etiological explanation of SM and, ultimately, SAD. In order to do so, MEAR and phonophobia data were collected from participants recruited from Bard College. Participants completed an anxiety measure before and after reading words aloud (which produced Δ STAI scores), an S/N threshold detection measure (which produced WIN scores), and an ART detection measure. SAM ratings and dB Volume levels were also collected as internal validity checks.

Broadly, it was hypothesized that impaired MEAR functioning would be positively correlated with phonophobia, as impaired MEAR functioning is associated with both high sympathetic activity and an inability to filter out one's own voice (Stephen W. Porges & Lewis, 2010); specifically, it was hypothesized that there would be a positive correlation between WIN scores and Δ STAI scores, there would be a positive correlation between ARTs and Δ STAI scores, and these two correlations would be comparable. Additionally, WIN scores, ARTs, SAM ratings, and initial (pre-reading activity) STAI scores were predicted to reflect normative data for

each of these measures, and dB Volume levels were predicted to be the same within the high Δ STAI scores and low Δ STAI scores groups.

The hypothesis that impaired MEAR functioning would be positively correlated with phonophobia was not supported. No association was found between WIN scores and Δ STAI scores in either ear, or between ARTs and Δ STAI scores in the left ear. Comparisons between these two correlations also yielded no significance. When a median split of Δ STAI scores was performed, no differences were found between ARTs or WIN scores in the high Δ STAI scores group and in the low Δ STAI scores group in either ear.

While a significant, positive correlation was found between ARTs and Δ STAI scores in the right ear ($p = .04$), one should exercise caution in interpreting these results. Data were reported for only the left ear in part to mitigate the possibility of a Type I error. If this right ear *p*-value were corrected for multiple comparisons, it would likely not remain significant. Evidence of this is the clear lack of significance of the left ear ART and Δ STAI score correlation ($p = .55$), despite no difference in means of left ear ART and right ear ART and no difference in correlations between left ear ART and Δ STAI and right ear ART and Δ STAI. This lack of a difference between left ear ART data and right ear ART data is in line with prior literature, which typically shows an agreement of ART data between ears (e.g., da Cruz Fernandes, Momensohn-Santos, Martins Carvalho, & de Queiroz Carvalho, 2013; Day & Feeney, 2008; Gates, Cooper, Kannel, & Miller, 1990). It would therefore be preemptive to draw conclusions based on this sole significant ART and Δ STAI correlation.

The significant ART and Δ STAI correlation in the right ear is all the more unconvincing given the lack of variation in right ear ART readings: only one participant had a right ear ART of 100 dB, meaning the remaining 27 participants had ARTs of either 85, 90, or 95 dB. Although

the Spearman's r_s correlation performed was corrected for data ties, this variance may simply be too narrow to make claims of significance. For this reason, a median split of Δ STAI scores was used to compare right ear ARTs in the high Δ STAI scores group with those in the low Δ STAI scores group. This analysis yielded no significant results, which makes the right ear ART and Δ STAI correlation all the more questionable. Furthermore, researchers investigating ARTs among people with SM similarly found greater significance with right ear data than with left ear data; however, when they replicated their study with a larger sample size, this effect disappeared (Henkin & Bar-Haim, 2015). All in all, this suggests that the significant correlation found between Δ STAI and right ear ART is not sufficient to support the hypothesis that impaired MEAR functioning would be positively correlated with phonophobia.

As predicted, WIN scores, ARTs, SAM ratings, and initial (pre-reading activity) STAI scores were in line with normative data. This means the lack of support of the main hypothesis likely cannot be attributed to unusual patterns among participants in the present study. Decibel levels were also as expected: there was no difference between dB Volume levels within the high Δ STAI scores group and dB Volume levels within the low Δ STAI scores group. Participants were therefore probably not compensating for phonophobia by speaking quietly, as doing so would have been expected to manifest in lower dB Volume levels within the low Δ STAI scores group.

Reflections on Methodology

In interpreting these results, it is necessary to determine whether they provide sufficient evidence to discredit the proposed etiological model (see Figure 2). One possibility is that the proposed model still holds, but limitations in the study's design account for the lack of support of the main hypothesis. Perhaps the most convincing argument for this possibility is the fact that

WIN score was not found to correlate with ART. This contradicts prior findings that S/N threshold correlates with ART among nonclinical and clinical populations, including people with autism, auditory processing disorders, Bell's palsy, and Williams syndrome (e.g., Harkrider & Smith, 2005; Henkin & Bar-Haim, 2015; Porges et al., 2013). Given the use of a nonclinical population in the present study, these results were expected to be replicated. The fact that they were not could suggest a methodological flaw in the present study.

One possible avenue for such an error is the measurement of ARTs. I am very grateful to have been lent the tympanometer, but due to time and budget constraints, some aspects of it might not have been ideal for my specific purposes. For instance, all measurements were made ipsilaterally, meaning the stimulus tone and measurement probe were presented on the same side, as opposed to contralaterally, which would involve presenting the stimulus tone to one ear while measuring the ART of the other ear. Some studies have found that contralateral readings are less sensitive to artifacts than are ipsilateral readings, as an ipsilateral measurement probe could inadvertently record the impedence of the auditory canal in addition to that of the eardrum (Kunov, 1977). However, studies have found that ipsilateral and contralateral readings are usually comparable, and if they are not, ipsilateral readings are affected uniformly (Harkrider & Smith, 2005); in fact, researchers have been unable to replicate these artifact findings, and have even found ipsilateral readings to be less prone to artifact, as these reflex pathways are more direct and independent, in that they are not affected by abnormalities in the other ear (Rawool, 2001). In other words, it is unlikely that there was an artifact effect resulting from the ipsilateral method, but if there was, it likely affected all participants in the same way, and thus should not have impacted comparisons.

If anything, ipsilateral ART testing may have been better suited to the present study than contralateral ART testing. The differences between the ARTs of participants with SM and those of nonclinical control participants are slightly more significant among ipsilateral readings than they are among contralateral readings (Bar-Haim et al., 2004; Henkin & Bar-Haim, 2015), meaning ipsilateral readings provide a more reliable construct for making inferences that are potentially generalizable to people with SM. This could be due to the fact that ipsilateral ART readings are more tightly correlated with S/N thresholds than are contralateral ART readings (Harkrider & Smith, 2005). Perhaps this is why audiologists report that supplementing ipsilateral ART readings with contralateral ART readings has become decreasingly popular (Emanuel, Henson, & Knapp, 2012; L. A. Perry, personal communication, November 19, 2015). All in all, being limited to ipsilateral ART readings was unlikely to have negatively affected results; that said, future studies should consider including contralateral readings to ensure that this is the case.

Although the belief that ipsilateral ART readings are more susceptible to artifact resulting from probe tone has largely been discredited, debate about appropriate probe tone has remained. The standard for probe tones is 226 Hz, but some researchers have investigated whether higher frequencies or broadband measurements (wherein the probe tone sweeps across multiple frequencies, rather than just one) are more accurate (Hunter & Shahnaz, 2014). The tympanometer used in the present study emitted a probe tone of 226 Hz, so this could potentially influence results. However, while results have been mixed, the consensus is that 226 Hz is an adequate frequency for detecting ARTs among humans who are at least six months old (Hunter & Shahnaz, 2014). It is therefore unlikely that probe tone impacted results in the present study.

An aspect of the tympanometer that may have had a substantial effect on results is the gradation of the measurements. ARTs were measured in 5 dB steps spanning from 85 to 100 dB,

which means each ART could be one of four possible discrete measurements (not including readings of "no response," which were excluded). These steps may have not been finely graded enough for a correlational analysis, as details that could be integral to interpretation may have been missing. In addition, the left ear ART frequency distribution of data from the present study appears to be truncated. This may be because all possible measurements are well within the range of normal ARTs (Hunter & Shahnaz, 2014). Higher measurements were constricted by an adherence to safe noise exposure levels, which cannot (and should not) be avoided, but the lowest measurement, 85 dB, is unusually high. Hunter and Shahnaz (2014) recommend starting measurements at a maximum of 80 dB, and most studies begin testing for ARTs at 70 dB (e.g., Bar-Haim et al., 2004; Laukli & Mair, 1980). This is likely why the distribution of ARTs in the present study appears truncated, while studies that test a wider range of measurements found normal distributions (e.g., Laukli & Mair, 1980). It is probable that some participants at the 85 dB level and perhaps the 100 dB level had ARTs outside the examined range. This is evidenced by the fact that some participants with 85 dB ARTs had ARTs detected at the 85 dB level on all six trials (three per ear), while others showed more variability; those with more variable ARTs probably truly had ARTs of 85 dB, while those whose MEARs were detected at 85 dB on all trials likely had lower ARTs that were not captured. The arbitrary binning of ART measurements may have masked variability that would have been useful for analysis. In future studies, tympanometer data should be collected using a wider range of dB and more finely graded steps.

Another consequence of the narrow range of available ART measurements is that there is no differentiation among abnormal ARTs. For instance, one study found that 87% of participants with SM have abnormal ARTs (defined as >100 dB)[4]. Of these 87% of participants, 85% had no

[4] For context, there were two such cases in the present study, each affecting only one ear.

ART even at the highest stimulus level, which was 120 dB. The 13% of participants with SM who did have normal ARTs still had significantly higher ARTs than did participants without SM (Bar-Haim et al., 2004). Although most participants with SM had abnormal ARTs, researchers were forced to split them into normal versus abnormal ART categories for analysis. It remains unclear whether participants had extremely high ARTs or entirely absent ARTs, as well as whether the distribution of ARTs among people with SM is normal. Because it is impossible to determine this without inflicting serious hearing damage, we are left guessing whether people with SM exhibit normal variation in ARTs with a much higher mean than that of a nonclinical population, or have qualitatively different MEARs. Although I would expect weakened vagal activity to predict a normal distribution of extremely high ARTs, it is impossible to test this using current tympanometric technology. Future studies should explore possible methods for gauging abnormal ARTs without causing hearing damage.

This lack of clarity highlights a major methodological problem: if people with SM have qualitatively different MEARs, it might not be reasonable to expect findings from a nonclinical population to be generalizable to those with SM. It is important to note that this study specifically focused on an auditory component of SM so as to elucidate the contribution of this particular mechanism to the disorder; that being said, it is not necessarily appropriate to examine this mechanism in isolation. Various aspects of SM might interact to create the particular set of symptoms characteristic of the disorder in such a way that is not compatible with nonclinical research. For instance, vagal dysregulation among those with SM could potentially affect laryngeal and pharyngeal recruitment, which could be a necessary component of the development of phonophobia. While it would be preemptive to discredit the possibility that the present findings are sufficiently generalizable to inform our understanding of SM, assuming this

possibility is the case would be equally unwarranted. Future research is necessary to determine whether results from the present study are representative of people with SM.

One concern regarding the generalizability of the present study to those with SM is the nature of participant recruitment. Because participation in research involves voluntarily entering a foreign, one-on-one situation with a stranger, it could be that participants underrepresented the prevalence of social anxiety in the U.S. college-attending population. Given the high comorbidity of SM and SAD, this means participants could have different physiological mechanisms from people with SM, which would compromise the applicability of the study to people with SM. Additionally, since I am a student at Bard and therefore know some students better than others, the study sample was in part a sample of convenience. This decreases the chances of obtaining a random sample, as people who knew me were more likely to participate. That said, the fact that I was the only experimenter mitigates the potential confound of participants being treated wholly differently. Likewise, all participants were run in the same location, so environmental variables were minimized. Still, if people with SM have fundamentally different MEARs, the possible underrepresentation of social anxiety could be a major consideration.

Conversely, it is equally possible that social anxiety was accurately represented, or even overrepresented. Bard is a small liberal arts school, which could potentially be more attractive than a large school to people who tend to experience more social anxiety. If this is the case, social anxiety could have been a possible confound during the reading task. Although I did take measures to mitigate performance anxiety, such as explicitly not paying attention during the reading task, telling participants not to worry about pronunciation, and forewarning them that I would be monitoring the dB Volume application, rather than listening to them speak, my

presence may have had an effect. If some participants still experienced anxiety resulting from my presence, though, this would have served to make the sample more closely resemble people with SM.

The possibility of anxiety being overrepresented in the sample could have also resulted in lower Δ STAI scores: if participants were still acclimating to the novelty of the experiment when they filled out the initial STAI, they could have become more comfortable by the time they completed the post-reading activity STAI. This would prevent the initial STAI from accurately measuring baseline anxiety, and could potentially mask any increases in anxiety that result from the reading aloud task. I would not expect this to be the case, though, as Holzman, Rousey, and Snyder (1966) recommend an adjustment period of about 30 seconds for a novel task. Going over the consent form alone easily surpassed this suggested timeframe. It is therefore unlikely that anxiety resulting from the novelty of the task would have impacted Δ STAI scores. If anything, physiological responses to the sound of one's own voice may have died down by the time the post-reading activity STAI was filled out, as participants spoke for about 2 minutes and 13 seconds. Still, since participants continuously heard their own voices throughout this time, I would expect reactions to have persisted for the duration of the task. Additionally, participants filled out the post-reading activity STAI directly after the reading activity so that their reactions to hearing their own voices would be fresh. All in all, it is unlikely that the timing of the task created confounds.

Although the timing of the task was expected to be appropriate, this and other aspects of the reading activity's validity should be examined in the future. The STAI is a reliable and valid measurement of transient, state-dependent anxiety, and ANEW ratings have been demonstrated to accurately gauge word valence (Soares et al., 2012). Furthermore, participants did actually

rate the words as neutral, meaning the present study replicated these findings. In fact, when rating word valence, many participants reported struggling to remember the words they had read a moment prior. Because words most likely to remain salient in memory are those with particularly positive or particularly negative valence (e.g., Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), this anecdotal evidence supports the assertion that participants were likely not distracted by word content. Initial STAI scores were also in accordance with normative data. This suggests that each component of the reading task was effective, at which point the task itself might not have captured the intended construct.

In the future, studies should be conducted to evaluate the construct validity of the reading task as a measure of phonophobia. Because I was unable to find an existing measure of phonophobia, this evaluation cannot be as straightforward as comparing the reading task to a well-validated, pre-existing measure. Instead, researchers might consider comparing brain activity elicited by this task with brain activity during self-reported phonophobia, or perhaps testing the correlation between data from this task and data from an implicit-association test that measures the association between one's own voice and anxiety or fear (Greenwald, McGhee, & Schwartz, 1998). If such comparisons are found to be significant, this would support the construct validity of the reading task, which would in turn support the present findings; if the validity of the reading task is not supported, this would indicate a methodological flaw that would call the present findings into question.

Compensatory Mechanism

Although validity of the reading activity cannot be assumed, data from all other measures were found to reflect normative data. This supports the possibility that the present findings were not influenced by methodological flaws, and that the proposed model is thus inaccurate. At this point, it is necessary to consider other explanations for the results. One such explanation is that people with SM could have a mechanism to compensate for excess auditory information. A study found that children with SM had very mature ventrolateral prefrontal cortices (VLPFCs) for their age, which allowed them to suppress irrelevant sounds that were inadequately filtered due to MEAR dysfunction (Henkin, Feinholz, Arie, & Bar-Haim, 2010). If this compensatory mechanism is not unique to those with SM, participants in the present study with particularly high ARTs may have similarly countered excess auditory information with VLPFC activity; this could explain the lack of correlation between ARTs and WIN scores.

If people with SM are able to compensate for MEAR dysfunction, it is necessary to reevaluate how their inability to speak emerges. One possibility lies in the function of the VLPFC. In addition to aiding in sensory gating (i.e., suppressing irrelevant sensory information), the VLPFC is a domain general inhibition center that filters out memories, words, and actions that are inappropriate for the task at hand (Anderson & Weaver, 2009). Among adults, these functions generally occur within the right hemisphere. However, children who are especially proficient at inhibiting irrelevant information show activation in the VLPFC within the left hemisphere; this lateralization is thought to reverse as the prefrontal cortex develops with age (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002). The left VLPFC contains Broca's area, which is implicated in word selection as well as language retrieval and production, specifically by inhibiting competing words (Anderson & Spellman, 1995; Buckner, Raichle, &

Petersen, 1995; Dronkers, Redfern, & Knight, 2000). Because children with SM most likely perform both word selecting functions and auditory gating functions within the left VLPFC, the substantial activity required of the left VLPFC to filter out irrelevant sounds could also work to inhibit language selection. Although the VLPFC generally inhibits only inappropriate words and stimuli, when it is stimulated externally (e.g., using transcranial magnetic stimulation), it inhibits even appropriate information (Anderson & Weaver, 2009). Sensory gating activity within the left VLPFC could therefore overstimulate language suppression, which could create the perfect storm for the development of SM. Conversely, left VLPFC activity could be occupied by the need for sensory gating, and thus be unable to engage in language retrieval and sensory gating simultaneously. Transcranial magnetic stimulation could be utilized in future studies to investigate the effects of VLPFC stimulation and inhibition on language retrieval among those with SM.

This explanation could account for the finding that people with SM have particular difficulty simultaneously attending to auditory information and vocalizing (Henkin & Bar-Haim, 2015). It could also explain the fact that people tend to grow out of SM during adolescence (American Psychiatric Association, 2013): as their prefrontal cortices develop, they likely switch to using the right VLPFC to suppress excess auditory information but continue to activate the left VLPFC during word selection, as is typical of adults (Bunge et al., 2002). When these processes occur in different hemispheres, sensory gating would no longer necessarily recruit inhibition of language access, or take up attentional and metabolic resources that would otherwise go toward language retrieval.

This alternative explanation could also help explain how SAD develops. The two instances when people with SM are most likely to be unable to speak are in crowds and with

strangers (American Psychiatric Association, 2013). These also happen to be the instances that are the least conducive to aural comprehension. In a crowd, there is a lot of excess background noise, and strangers have unfamiliar speech patterns. Both of these occurrences would normally elicit the MEAR, but as people with SM have dampened MEARs, they would likely rely on the VLPFC to compensate, and thus have difficulty speaking. This could potentially be anxiety provoking, and because vagal dysregulation would prevent physiological calming, it would be difficult to shift out of this anxious state. In turn, the MEAR would continue to be underactive, meaning the compensatory mechanism would still be necessary, so the inability to speak would continue or worsen. People with SM might then feel afraid or pressured to speak in social situations, which would increase the likelihood of sympathetic arousal when entering social situations, and thus cause this cycle to repeat. Eventually, this could contribute to a building association of social situations with anxiety, leading to the development of SAD.

This compensatory mechanism as well as the original proposed model should be explored in future research. If the compensatory mechanism is investigated, researchers should take care to ensure that the patterns of activation typical of children with mature VLPFCs (i.e., left lateralization) is mirrored in children with SM. This can be accomplished using functional imaging and/or by examining volumetric properties of left VLPFCs among children with SM. In order to corroborate the finding that VLPFC activity serves a compensatory function, its relationship to ARTs should be investigated among people with SM. If the VLPFC does counteract unfiltered auditory information, I would expect people with the highest ARTs to show the most VLPFC activation. VLPFC activity, as well as general brain activity, should also be monitored while people with SM process incoming vocalizations, preferably both while exhibiting SM symptoms and while not exhibiting symptoms. This could help determine what

happens differently in situations that elicit mutism from those that do not, as well as how integral a role abnormal auditory efferent activity plays in symptom presentation. Additionally, these findings could clarify the feasibility of the original proposed model.

Conclusion

Results from the present study do not support the hypothesis that impaired MEAR functioning, as determined by ART or WIN score, would be positively correlated with phonophobia. However, whether this is because of methodological factors, because the mechanisms of action are different within a clinical population, or because of a flaw in the proposed model remains to be determined. Each of these explanations presents avenues for future exploration that could improve the current understanding of SM.

This study is the first in the literature to integrate recent findings regarding vagal and audiological abnormalities among people with SM. In addition, it proposed a novel method for measuring phonophobia. Taken together, these advances will hopefully serve as jumping off points for future thought. Compared to the current basis for treatment of SM, which is entirely contingent upon the false belief that SM is a severe form of SAD, this study takes great strides. Recommended future directions will hopefully continue to clarify the mechanisms that underlie SM, even if this is through a process of elimination. Additionally, if the present study is replicated with a clinical population, findings could determine whether people with SM have qualitatively different auditory efferent systems. This information could aid our understanding of SM as well as our understanding of the auditory efferent system generally. The more we can glean about SM, the closer we get to developing successful treatment models.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100.

Anderson, M. C., & Weaver, C. (2009). Inhibitory control over action and memory. In L. R. Squire (Ed.), *Encyclopedia of neuroscience* (5th ed., pp. 153–163). Oxford, UK: Academic Press.

Aron, A., Coups, E. J., & Aron, E. N. (2013). *Statistics for psychology* (6th ed.). Upper Saddle River, NJ: Pearson.

Backs, R. W., da Silva, S. P., & Han, K. (2005). A comparison of younger and older adults' Self-Assessment Manikin ratings of affective pictures. *Experimental Aging Research*, *31*(4), 421–440. http://doi.org/10.1080/03610730500206808

Bar-Haim, Y., Henkin, Y., Ari-Even-Roth, D., Tetin-Schneider, S., Hildesheimer, M., & Muchnik, C. (2004). Reduced auditory efferent activity in childhood selective mutism. *Biological Psychiatry*, *55*(11), 1061–1068. http://doi.org/10.1016/j.biopsych.2004.02.021

Bayrampour, H., McDonald, S., Fung, T., & Tough, S. (2014). Reliability and validity of three shortened versions of the State Anxiety Inventory scale during the perinatal period. *Journal of Psychosomatic Obstetrics and Gynecology*, *35*(3), 101–107. http://doi.org/10.3109/0167482X.2014.950218

Black, B., & Uhde, T. W. (1992). Elective mutism as a variant of social phobia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(6), 1090–1094. Retrieved from http://ezproxy.usherbrooke.ca/login?url=https://search.ebscohost.com/login.aspx?direct=tru

e&db=mnh&AN=1342579&site=ehost-live

Black, B., & Uhde, T. W. (1995). Psychiatric characteristics of children with selective mutism: A pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*(7), 847–856. http://doi.org/10.1097/00004583-199507000-00007

Blumberg, B., Cooper, D. R., & Schindler, P. S. (2011). Measures of association. In *Business research methods* (3rd ed., pp. 589–616). New York, NY: McGraw-Hill Higher Education.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7962581

Bradley, M. M., & Lang, P. J. (1999). Affective Norms for English Words (ANEW): Instruction manual and affective ratings. *Psychology*, *Technical*(C-1). http://doi.org/10.1109/MIC.2008.114

Brodal, P. (2004). *The central nervous system: Structure and function* (3rd ed.). New York, NY: Oxford University Press. Retrieved from http://myaccess.library.utoronto.ca/login?url=http://search.proquest.com/docview/622170471?accountid=14771\nhttp://bf4dv7zn3u.search.serialssolutions.com/?ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/PsycINFO&rft_val_fmt=info:ofi/fmt:kev:

Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *Journal of Neurophysiology*, *74*(5), 2163–2173. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8592204

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002).

Immature frontal lobe contributions to cognitive control in children. *Neuron*, *33*(2), 301–311. http://doi.org/10.1016/S0896-6273(01)00583-9

Carlson, J. S., Kratochwill, T. R., & Johnston, H. (1994). Prevalence and treatment of selective mutism in clinical practice: A survey of child and adolescent psychiatrists. *Journal of Child and Adolescent Psychopharmacology*, *4*(4), 281–291. http://doi.org/10.1089/cap.1994.4.281

Carlson, J. S., Mitchell, A. D., & Segool, N. (2008). The current state of empirical support for the pharmacological treatment of selective mutism. *School Psychology Quarterly*, *23*(3), 354–372. http://doi.org/10.1037/1045-3830.23.3.354

Chlan, L., Savik, K., & Weinert, C. (2003). Development of a shortened state anxiety scale from the Spielberger State-Trait Anxiety Inventory (STAI) for patients receiving mechanical ventilatory support. *Journal of Nursing Measurement*, *11*(3), 283–293. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15633782

Consonni, E. B., Calderon, I. M., Consonni, M., De Conti, M. H., Prevedel, T. T., & Rudge, M. V. (2010). A multidisciplinary program of preparation for childbirth and motherhood: Maternal anxiety and perinatal outcomes. *Reproductive Health*, *7*, 28. http://doi.org/10.1186/1742-4755-7-28

Court, H., Greenland, K., & Margrain, T. H. (2010). Measuring patient anxiety in primary care: Rasch analysis of the 6-item Spielberger state anxiety scale. *Value in Health*, *13*(6), 813–819. http://doi.org/10.1111/j.1524-4733.2010.00758.x

Crumley, F. E. (1990). The masquerade of mutism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *29*(2), 319.

da Cruz Fernandes, L., Momensohn-Santos, T. M., Martins Carvalho, J. S., & de Queiroz Carvalho, F. L. (2013). Tinnitus and normal hearing: A study on contralateral acoustic

reflex. *American Journal of Audiology*, *22*(2), 291–296. http://doi.org/10.1044/1059-0889(2013/13-0005)

Day, J. E., & Feeney, M. P. (2008). The effect of the 226-Hz probe level on contralateral acoustic stapedius reflex thresholds. *Journal of Speech, Language, and Hearing Research*, *51*(4), 1016–1025. http://doi.org/10.1044/1092-4388(2008/074)

dB Volume (Version 2.4) [iPhone application]. (2016). Germany: DSP Mobile.

de Andrade, K. C. L., Neto, S. C., & Menezes, P. de L. (2011). The importance of acoustic reflex in speech discrimination. In I. Ipsic (Ed.), *Speech Technologies*. InTech. Retrieved from http://www.intechopen.com/books/speech-technologies/the-importance-of-acoustic-reflex-in-speech-discrimination

Dronkers, N. F., Redfern, B. B., & Knight, R. T. (2000). The neural architecture of language disorders. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 949–958). Cambridge, MA: The MIT Press.

Dummit, E. S., Klein, R. G., Tancer, N. K., Asche, B., Martin, J., & Fairbanks, J. a. (1997). Systematic assessment of 50 children with selective mutism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(5), 653–660. http://doi.org/10.1097/00004583-199705000-00016

Emanuel, D. C., Henson, O. E. C., & Knapp, R. R. (2012). Survey of audiological immittance practices. *American Journal of Audiology*, *21*(1), 60–75. http://doi.org/10.1044/1059-0889(2012/11-0037)

Ferman, L., Verschuure, J., & Van Zanten, B. (1993). Impaired speech perception in noise in patients with a normal audiogram. *Audiology*, *32*(1), 49–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8447761

Ferré, P., Guasch, M., Moldovan, C., & Sánchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods*, *44*(2), 395–403. http://doi.org/10.3758/s13428-011-0165-x

Fishbach, A., Eyal, T., & Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass*, *4*(8), 517–530. http://doi.org/10.1111/j.1751-9004.2010.00285.x

Fishbach, A., & Finkelstein, S. R. (2011). How feedback influences persistence, disengagement, and change in goal pursuit. In H. Aarts & A. J. Elliot (Eds.), *Goal-directed behavior* (pp. 203–230). New York, NY: Psychology Press.

Garcia, A. M., Freeman, J. B., Francis, G., Miller, L. M., & Leonard, H. L. (2004). Selective Mutism. In T. H. Ollendick & J. S. March (Eds.), *Phobic and anxiety disorders in children and adolescents: A clinician's guide to effective psychosocial and pharmacological interventions* (pp. 433–455). New York, NY: Oxford University Press.

Gates, G. A., Cooper, J. C., Kannel, W. B., & Miller, N. J. (1990). Hearing in the elderly: the Framingham cohort, 1983-1985. Part I. Basic audiometric test results. *Ear and Hearing*, *11*(4), 247–256. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2210098

Gershon, R. C., Wagster, M. V, Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH Toolbox for Assessment of Neurological and Behavioral Function. *Neurology*, *80*(11), S2–S6. http://doi.org/10.1212/WNL.0b013e3182872e5f

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. http://doi.org/10.1037/0022-3514.74.6.1464

Haahr, M. (n.d.). List randomizer. Retrieved January 1, 2016, from

https://www.random.org/lists/

Harkrider, A. W., & Smith, S. B. (2005). Acceptable noise level, phoneme recognition in noise, and measures of auditory efferent activity. *Journal of the American Academy of Audiology*, *16*(8), 530–545. http://doi.org/10.3766/jaaa.16.8.2

Heilman, K. J., Connolly, S. D., Padilla, W. O., Wrzosek, M. I., Graczyk, P. a., & Porges, S. W. (2012). Sluggish vagal brake reactivity to physical exercise challenge in children with selective mutism. *Development and Psychopathology*, *24*(01), 241–250. http://doi.org/10.1017/S0954579411000800

Henkin, Y., & Bar-Haim, Y. (2015). An auditory-neuroscience perspective on the development of selective mutism. *Developmental Cognitive Neuroscience*, *12*, 86–93. http://doi.org/10.1016/j.dcn.2015.01.002

Henkin, Y., Feinholz, M., Arie, M., & Bar-Haim, Y. (2010). P50 suppression in children with selective mutism: A preliminary report. *Journal of Abnormal Child Psychology*, *38*(1), 43–48. http://doi.org/10.1007/s10802-009-9346-9

Hirsch, C. R., & Clark, D. M. (2007). Imagery special issue: Underestimation of auditory performance in social phobia and the use of audio feedback. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(4), 447–458. http://doi.org/10.1016/j.jbtep.2007.08.004

Ho, S. M. Y., Mak, C. W. Y., Yeung, D., Duan, W., Tang, S., Yeung, J. C., & Ching, R. (2015). Emotional valence, arousal, and threat ratings of 160 Chinese words among adolescents. *PLoS ONE*, *10*(7), 1–13. http://doi.org/10.1371/journal.pone.0132294

Holzman, P. S., & Rousey, C. (1966). The voice as a percept. *Journal of Personality and Social Psychology*, *4*(1), 79–86. http://doi.org/10.1037/h0023518

Holzman, P. S., Rousey, C., & Snyder, C. (1966). On listening to one's own voice: Effects on

psychophysiological responses and free associations. *Journal of Personality and Social Psychology*, *4*(4), 432–441. http://doi.org/10.1037/h0023790

Hunter, L. L., & Shahnaz, N. (2014). *Acoustic immittance measures: Basic and advanced practice*. (B. Stach, Ed.). San Diego, CA: Plural Publishing Inc.

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology*, *25*(4), 652–665. http://doi.org/10.1016/j.jcps.2014.12.002

Interacoustics. (2015, November 30). *Acoustic reflexes: An introduction*. Retrieved from https://www.youtube.com/watch?v=3a3Eeuhkh-c

Jalenques, I., Enjolras, J., & Izaute, M. (2013). Valence émotionnelle des mots dans la schizophrénie. *L'Encéphale: Revue de Psychiatrie Clinique Biologique et Thérapeutique*, *39*(3), 189–197. http://doi.org/10.1016/j.encep.2012.06.011

Jeger, A. M., & Goldfried, M. R. (1976). A comparison of situation tests of speech anxiety. *Behavior Therapy*, *7*(2), 252–255. http://doi.org/10.1016/S0005-7894(76)80284-5

Kallen, M., Slotkin, J., Griffith, J., Magasi, S., Salsman, J., Nowinski, C., & Gershon, R. (2012). NIH Toolbox Technical Manual. National Institutes of Health and Northwestern University. Retrieved from http://www.nihtoolbox.org/HowDoI/TechnicalManual/Technical Manual sections/Toolbox Words_in_Noise Test Technical Manual.pdf

Keene, K., Merovitz, A., Irvine, E., Manji, N., Everett, M., Chung, I., … Chan, B. (2013). Accuracy of smartphone sound level meter applications. *Canadian Hearing Report*, *8*(6), 24–28.

Kogan, A., Oveis, C., Carr, E., & Gruber, J. (2014). Vagal activity is quadratically related to prosocial traits, prosocial emotions, and observer perceptions of prosociality. *Journal of*

Personality and Social Psychology, *107*(6), 1051–1063. Retrieved from http://psycnet.apa.org/psycinfo/2014-39093-001/

Kunov, H. (1977). The "eardrum artifact" in ipsilateral reflex measurements. *Scandinavian Audiology*, *6*(3), 163–166. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/929094

Lamb, D. H. (1972). Speech anxiety: Towards a theoretical conceptualization and preliminary scale development. *Speech Monographs*, *39*(1), 62–67. http://doi.org/10.1080/03637757209375739

Laukli, E., & Mair, I. W. S. (1980). Ipsilateral and contralateral acoustic reflex thresholds. *Audiology*, *19*(6), 469–479. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7425952

Lent, R. W., Russell, R. K., & Zamostny, K. P. (1981). Comparison of cue-controlled desensitization, rational restructuring, and a credible placebo in the treatment of speech anxiety. *Journal of Consulting and Clinical Psychology*, *49*(4), 608–610. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6114970

Ludewig, K., Ludewig, S., Seitz, A., Obrist, M., Geyer, M. A., & Vollenweider, F. X. (2003). The acoustic startle reflex and its modulation: Effects of age and gender in humans. *Biological Psychology*, *63*(3), 311–323. http://doi.org/10.1016/S0301-0511(03)00074-7

Mack, A. (2011). e-Audiology. Retrieved from http://cs1.e-learningforhealthcare.org.uk/public/Audiology/Sessions/AUD_1_08_02/d/AUD_Session/373/session.html?lms=n

Maico Otowave 102 Hand Held Portable Tympanometer operating manual. (n.d.) (6th ed.). Eden Prairie, MN: Maico Diagnostics.

Manassis, K., Fung, D., Tannock, R., Sloman, L., Fiksenbaum, L., & McInnes, A. (2003). Characterizing selective mutism: Is it more than social anxiety? *Depression and Anxiety*,

18(3), 153–161. http://doi.org/10.1002/da.10125

Martini, F. H., Nath, J. L., & Bartholomew, E. F. (2015). *Fundamentals of Anatomy and Physiology* (10th ed.). San Francisco, CA: Pearson Education, Inc.

Mash, E. J., & Wolfe, D. A. (2013). *Abnormal child psychology* (5th ed.). Belmont, CA: Wadsworth.

McGuire, P. K., Silbersweig, D. a, & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*, 907–917. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8673501

Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172–175. http://doi.org/10.1037/0033-2909.111.1.172

Miller, N., Maruyama, G., Beaber, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, *34*(4), 615–624. http://doi.org/10.1037/0022-3514.34.4.615

Morris, J. D. (1995). SAM: the Self-Assessment Manikin. An efficient cross-cultural measurement of emotional response. *Journal of Advertising Research*, *35*(6), 63–68.

Muchnik, C., Ari-Even Roth, D., Hildesheimer, M., Arie, M., Bar-Haim, Y., & Henkin, Y. (2013). Abnormalities in auditory efferent activities in children with selective mutism. *Audiology and Neurotology*, *18*(6), 353–361. http://doi.org/10.1159/000354160

Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal : The Journal of Medical Association of Malawi*, *24*(3), 69–71.

Muris, P., & Ollendick, T. H. (2015). Children who are anxious in silence: A review on selective mutism, the new anxiety disorder in DSM-5. *Clinical Child and Family Psychology Review*,

151–169. http://doi.org/10.1007/s10567-015-0181-y

Nast, D. R., Speer, W. S., & Le Prell, C. G. (2014). Sound level measurements using smartphone "apps": Useful or inaccurate? *Noise and Health*, *16*(72), 251–256. http://doi.org/10.4103/1463-1741.140495

National Institutes of Health & Northeastern University. (2015). NIH Toolbox iPad App eLearning. Retrieved from http://www.nihtoolbox.org/HowDoI/HowToAdministerTheToolbox/e-LearningiPad/eLearning iPad app files/story_html5.html

NIH Toolbox Administrator's Manual. (2012). National Institutes of Health and Northwestern University. Retrieved from http://www.nihtoolbox.org/Resources/NIHToolboxiPadapp/NIH Toolbox iPad app documents/NIH Toolbox App Administrator%27s Manual v1 3 3-17-16.pdf

NIH Toolbox (Version 1.3) [iPad application]. (2015). National Institutes of Health and Northeastern University.

Niquette, P. A. (2012). Noise exposure: Explanation of OSHA and NIOSH safe exposure limits and the importance of noise dosimetry. *Canadian Hearing Report*, *9*(3), 22–29.

Nondahl, D. M., Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., & Dalton, D. S. (2013). 16-year change in acoustic admittance measures among older adults: Data from a population-based study. *Journal of Speech, Language, and Hearing Research*, *56*(6), 1745–1750. http://doi.org/10.1044/1092-4388(2013/12-0381)

Osinsky, R., Seeger, J., Mussel, P., & Hewig, J. (2016). Face-induced expectancies influence neural mechanisms of performance monitoring. *Cognitive, Affective and Behavioral Neuroscience*, *16*(2), 261–275. http://doi.org/10.3758/s13415-015-0387-y

Plack, C. J. (2014). *The sense of hearing* (2nd ed.). New York, NY: Psychology Press.

Plener, P., Gatz, S. A., Schuetz, C., Ludolph, A. G., & Kölch, M. (2012). A case of selective mutism in an eight-year-old girl with thalassaemia major after bone marrow transplantation. *Pharmacopsychiatry*, *45*(1), 37–39. http://doi.org/10.1016/S0924-9338(09)71189-7

Porges, S. W. (1995). Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. *Psychophysiology*, *32*, 301–318. http://doi.org/10.1111/j.1469-8986.1995.tb01213.x

Porges, S. W., Doussard-Roosevelt, J. A., Portales, A. L., & Greenspan, S. I. (1996). Infant regulation of the vagal "brake" predicts child behavior problems: A psychobiological model of social behavior. *Developmental Psychobiology*, *29*(8), 697–712. http://doi.org/10.1002/(SICI)1098-2302(199612)29:8<697::AID-DEV5>3.0.CO;2-O

Porges, S. W., & Lewis, G. F. (2010). The polyvagal hypothesis: Common mechanisms mediating autonomic regulation, vocalizations and listening. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalization: An integrative neuroscience approach* (1st ed., Vol. 19, pp. 255–264). London: Academic Press. http://doi.org/10.1016/B978-0-12-374593-4.00025-5

Porges, S. W., Macellaio, M., Stanfill, S. D., McCue, K., Lewis, G. F., Harden, E. R., … Heilman, K. J. (2013). Respiratory sinus arrhythmia and auditory processing in autism: Modifiable deficits of an integrated social engagement system? *International Journal of Psychophysiology*, *88*(3), 261–270. http://doi.org/10.1016/j.ijpsycho.2012.11.009

Rawool, V. W. (1998). Effect of probe frequency and gender on click-evoked ipsilateral acoustic reflex thresholds. *Acta Oto-Laryngologica*, *118*(3), 307–312. http://doi.org/10.1080/00016489850183377

Rawool, V. W. (2001). Can maximum comfortable loudness levels in hearing impaired listeners be predicted from ipsilateral acoustic reflex thresholds recorded with high frequency probes? *Scandinavian Audiology*, *30*(2), 96–105. http://doi.org/10.1080/010503901300112202

Slutsky, J. M., & Allen, G. J. (1978). Influence of contextual cues on the efficacy of desensitization and a credible placebo in alleviating public speaking anxiety. *Journal of Consulting and Clinical Psychology*, *46*(1), 119–125. http://doi.org/10.1037/0022-006X.46.1.119

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods*, *44*(1), 256–269. http://doi.org/10.3758/s13428-011-0131-7

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *State-Trait Anxiety Inventory for adults (Form Y)*. *Mind Garden*. Redwood City, CA: Mind Garden, Inc.

Spielberger, C. D., & Sydeman, S. J. (1994). State-trait anxiety inventory and state-trait anger expression inventory. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 292–321). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stadthagen-Gonzalez, H., Imbault, C., Pérez Sánchez, M. A., & Brysbaert, M. (2016). Norms of valence and arousal for 14,031 Spanish words. *Behavior Research Methods*. http://doi.org/10.3758/s13428-015-0700-2

Tluczek, A., Henriques, J. B., & Brown, R. L. (2009). Support for the reliability and validity of a six-item state anxiety scale derived from the State-Trait Anxiety Inventory. *Journal of*

Nursing Measurement, *17*(1), 19–28. http://doi.org/10.1891/1061-3749.17.1.19

van de Vijver, I., Ridderinkhof, K. R., & de Wit, S. (2015). Age-related changes in deterministic learning from positive versus negative performance feedback. *Neuropsychology, Development, and Cognition*, *22*(5), 595–619. http://doi.org/10.1080/13825585.2015.1020917

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. http://doi.org/10.3758/s13428-012-0314-x

Whitlock, M. C., & Schluter, D. (2008). *The analysis of biological data*. Greenwood Village, CO: Roberts and Company Publishers.

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology*, *14*(9), 453–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14708835

Wilson, R. H., & Burks, C. A. (2005). Use of 35 words for evaluation of hearing loss in signal-to-babble ratio: A clinic protocol. *Journal of Rehabilitation Research and Development*, *42*(6), 839–852. http://doi.org/10.1682/JRRD.2005.01.0009

Wilson, R. H., Carnell, C. S., & Cleghorn, A. L. (2007). The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *Journal of the American Academy of Audiology*, *18*(6), 522–529. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17849640

Wilson, R. H., & McArdle, R. (2007). Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) Test. *Journal of the American Academy of Audiology*, *18*(10), 813–825.

Yeganeh, R., Beidel, D. C., Turner, S. M., Pina, A. a, & Silverman, W. K. (2003). Clinical

distinctions between selective mutism and social phobia: An investigation of childhood psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(9), 1069–1075. http://doi.org/10.1097/01.CHI.0000070262.24125.23

Young, B. J., Bunnell, B. E., & Beidel, D. C. (2012). Evaluation of children with selective mutism and social phobia: A comparison of psychological and psychophysiological arousal. *Behavior Modification*, *36*(4), 525–544. http://doi.org/10.1177/0145445512443980

Zecker, S. G. (2012). Audition. In *Unveiling the NIH Toolbox*. Bethesda, MD. Retrieved from http://www.nihtoolbox.org/Publications/presentations/Pages/default.aspx

Zecker, S. G., Hoffman, H. J., Frisina, R., Dubno, J. R., Dhar, S., Wallhagen, M., … Wilson, R. H. (2013). Audition assessment using the NIH Toolbox. *Neurology*, *80*(11), S45–S48. http://doi.org/10.1212/WNL.0b013e3182872dd2

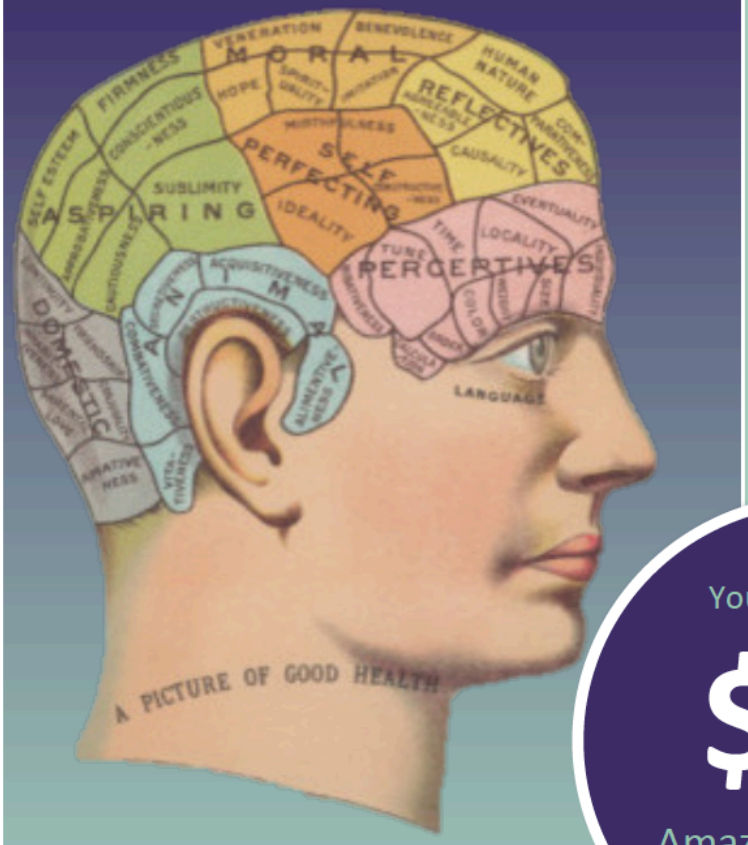
Zheng, Z. Z., MacDonald, E. N., Munhall, K. G., & Johnsrude, I. S. (2011). Perceiving a stranger's voice as being one's own: A "rubber voice" illusion? *PLoS ONE*, *6*(4), 1–8. http://doi.org/10.1371/journal.pone.0018655

Appendix A: Recruitment Flyer

Participate in a

Psych Study!

PI: Alexandra Batzdorf



Are you...

- **18 years or older?**
- **Fluent in English?**
- **Someone with normal hearing?**

Do you want FREE candy and a chance to win a $50 Amazon gift card? Then you qualify to participate in a short, simple 10-15 minute Psychology study that non-invasively investigates the middle ear and word processing. Visit http://bit.do/PsychStudy to learn about your own ears' reflexes, while contributing to scientific research! If you have any questions, please contact Alexandra Batzdorf at ab5673@bard.edu.

You could win a

$50

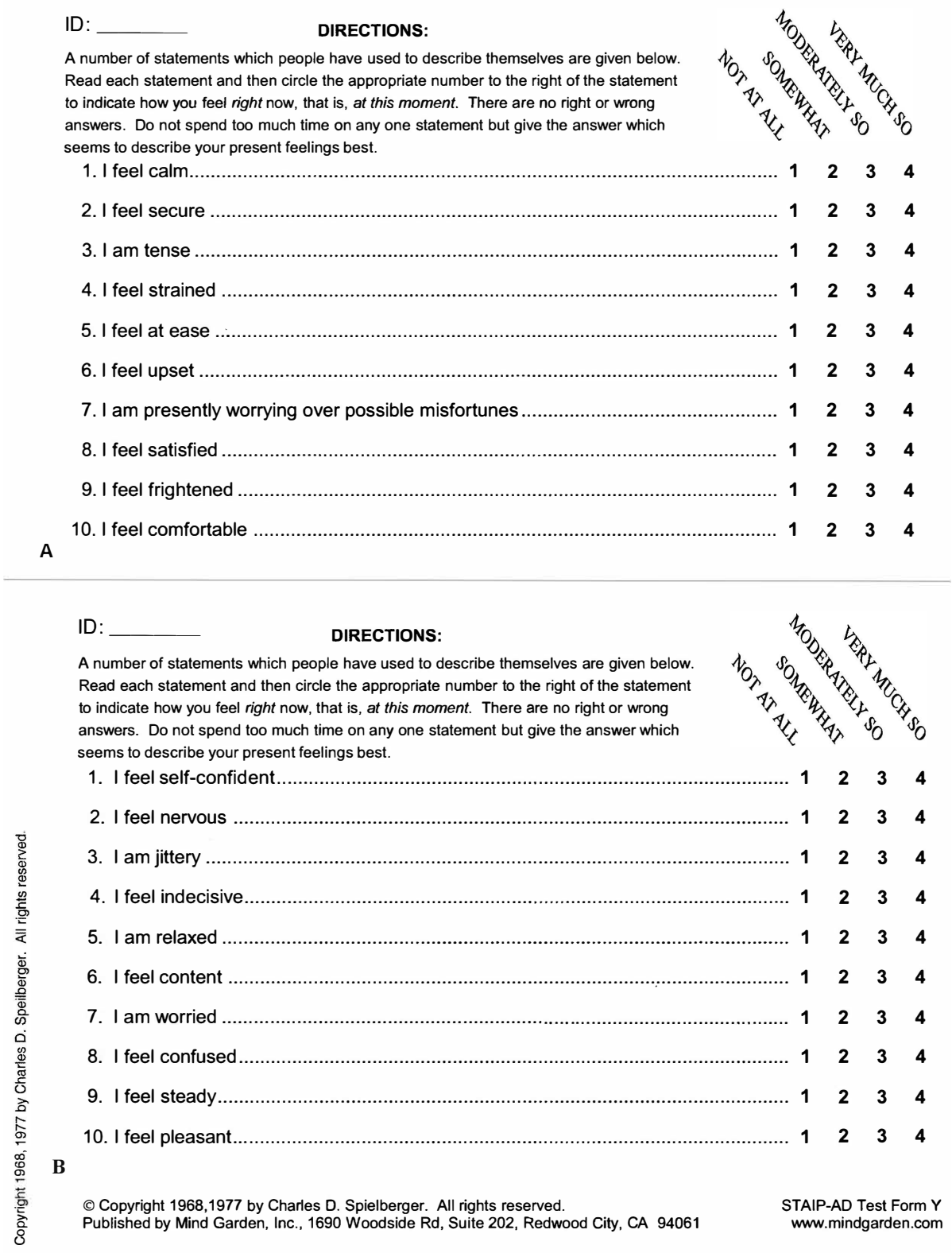
Amazon Gift Card!

Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy

Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy
Psychology Study Alexandra Batzdorf http://bit.do/PsychStudy

Appendix B: Measures

STAI



ID: _____

DIRECTIONS:

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	NOT AT ALL	SOMEWHAT	MODERATELY SO	VERY MUCH SO
1. I feel calm	1	2	3	4
2. I feel secure	1	2	3	4
3. I am tense	1	2	3	4
4. I feel strained	1	2	3	4
5. I feel at ease	1	2	3	4
6. I feel upset	1	2	3	4
7. I am presently worrying over possible misfortunes	1	2	3	4
8. I feel satisfied	1	2	3	4
9. I feel frightened	1	2	3	4
10. I feel comfortable	1	2	3	4

A

ID: _____

DIRECTIONS:

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	NOT AT ALL	SOMEWHAT	MODERATELY SO	VERY MUCH SO
1. I feel self-confident	1	2	3	4
2. I feel nervous	1	2	3	4
3. I am jittery	1	2	3	4
4. I feel indecisive	1	2	3	4
5. I am relaxed	1	2	3	4
6. I feel content	1	2	3	4
7. I am worried	1	2	3	4
8. I feel confused	1	2	3	4
9. I feel steady	1	2	3	4
10. I feel pleasant	1	2	3	4

B

STAIP-AD Test Form Y
www.mindgarden.com

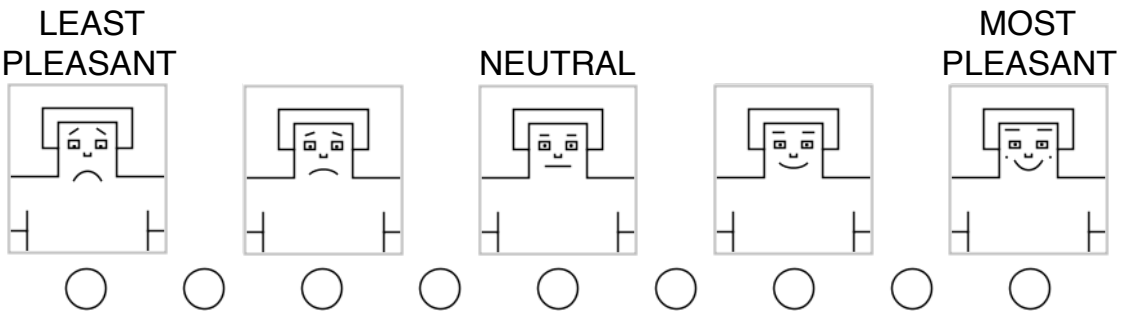
Neutral Words List

golfer	book	plant	hat	taxi	name	item
modest	street	icebox	corner	avenue	rain	farm
theory	activate	listless	cork	market	quart	coast
fabric	sphere	hairdryer	knot	mischief	violin	
hairpin	frog	subdued	patent	ink	journal	
bench	lamb	bowl	muddy	curtains	teacher	key
utensil	glass	salute	glacier	lawn	door	circle
lantern	banner	unit	village	poetry	square	barrel
passage	metal	ship	lamp	limber	bus	
lighthouse	industry	chair	ketchup	table	tower	
serious	slush	windmill	locker	skyscraper		
contents	method	humble	cabinet	kettle	office	
column	machine	museum	poster	seat	appliance	
moment	quiet	highway	material	patient	sentiment	
umbrella	part	prairie	trunk	vest	window	spray
reverent	odd	rock	headlight	cord	custom	
history	nonchalant		inhabitant	stiff	iron	
pamphlet	fur	context	thermometer		wagon	stove
basket	pencil	nonsense	kerchief	tool	corridor	
statue	errand	building	trumpet	truck	swamp	paper
month	mantel	nursery	paint	lightbulb		

SAM

ID: _____

Please take a moment to bubble in the most accurate rating of the overall **content of the words** you read aloud:



Alternative Valence Scale

ID: _____

Please take a moment to circle the number or dot that most accurately rates the overall **content of the words** you read aloud:

1	•	2	•	3	•	4	•	5
Very unpleasant		Somewhat unpleasant		Neither unpleasant nor pleasant		Somewhat pleasant		Very pleasant

Appendix C: Consent and Debriefing

Consent Form

You are being asked to participate in a research study conducted by Alexandra Batzdorf to investigate correlations between the middle ear acoustic reflex and word processing.

You will be asked to complete some short questionnaires, a simple reading activity, a listening activity that indirectly measures your acoustic reflex (ART), and a non-invasive test to measure the threshold needed to elicit ART. The acoustic reflex is characterized by the contraction of a muscle in your middle ear so as to filter out sudden low-frequency sounds.

What is the purpose of this study?

This study is being conducted to further scientific understanding of the connections between ART and word processing.

What is involved in this study?

The first phase of this study will require you to fill out a brief questionnaire about your comfort before and after reading a list of words aloud.

The second phase of this study will require you to listen to and repeat words that you will hear through a headset. Some of the words will be accompanied by background noise that sounds like being in a crowded room.

The third phase of this study will require you to sit still as a wand is placed against the outer opening of your ear canal. This may feel a bit like the ear buds of headphones, but unlike ear buds, the wand will NOT actually enter your ear. You will then hear a few tones at different frequencies and volumes.

How long will the study be?

The first phase will take about 10 minutes.

The second phase will take about 7 minutes.

The third phase should take under a minute.

What are the risks of the study?

There are no health risks associated with this study, but the questionnaires will require you to assess your comfort levels. This, as well as reading aloud, may cause some people mild discomfort. Additionally, while the ART wand does not pose any medical risk, it is a novel experience, so it may feel strange. You are free to leave and stop participating in this study at any time, for any reason. You will not be penalized if you do so, and you will still receive compensation.

If this study leads you to experience distress, please call BRAVE at (845) 758-7777 or Bard Counseling Services at (845) 758-7433. Feel free to ask me any questions at this time. If any questions or concerns about this study should arise at a later time, please contact the responsible investigator (me), Alexandra Batzdorf, at ab5673@bard.edu. If you have questions about your rights as a research participant, please contact the Bard Institutional Review Board at IRB@bard.edu.

Are there benefits to participating in this study?

Participating in this study will provide you with information about your own ears' physiological responses. The data you provide may further the scientific understanding and will aid me in my senior project endeavors.

Will I be compensated?

You will receive a piece of candy of your choosing and if you enter your email below, you will be entered to win a $50 Amazon gift card. **Your name and email will NOT be linked to the data you provide in this study.**

Whom can I contact about this study?

If you have questions or concerns about this study, please contact the responsible investigator, Alexandra Batzdorf, at ab5673@bard.edu. If you have questions about your rights as a research participant, please contact the Bard Institutional Review Board at IRB@bard.edu.

Feel free to ask any questions now, or at any point throughout the study.

STATEMENT OF CONSENT

By signing below, you are indicating that you are **at least 18 years old** and that you provide informed consent to participate in this study.

"The purpose of this study, procedures to be followed, risks and benefits have been explained to me. I have been given an opportunity to ask questions, and my questions have been answered to my satisfaction. I have been told whom to contact if I have additional questions. I have read this consent form and agree to be in this study, with the understanding that I may withdraw at any time."

Name: _____

Signature: _____

If you wish to be entered into the raffle to win a $50 Amazon gift card, please provide your email below. Remember, no identifying information will be connected to your data in any way. Your email address will only be used for the express purposes of contacting you should you win the gift card.

Email: _____

Debriefing Statement

Thank you for participating in this study. The purpose is to investigate the link between the acoustic reflex threshold (ART) and anxiety as a result of hearing your own voice. If this study leads you to experience distress, please call BRAVE at (845) 758-7777 or Bard Counseling Services at (845) 758-7433. Feel free to ask me any questions at this time. If any questions or concerns about this study should arise at a later time, please contact the responsible investigator (me), Alexandra Batzdorf, at ab5673@bard.edu. If you have questions about your rights as a research participant, please contact the Bard Institutional Review Board at IRB@bard.edu.

Appendix D: Right Ear Data

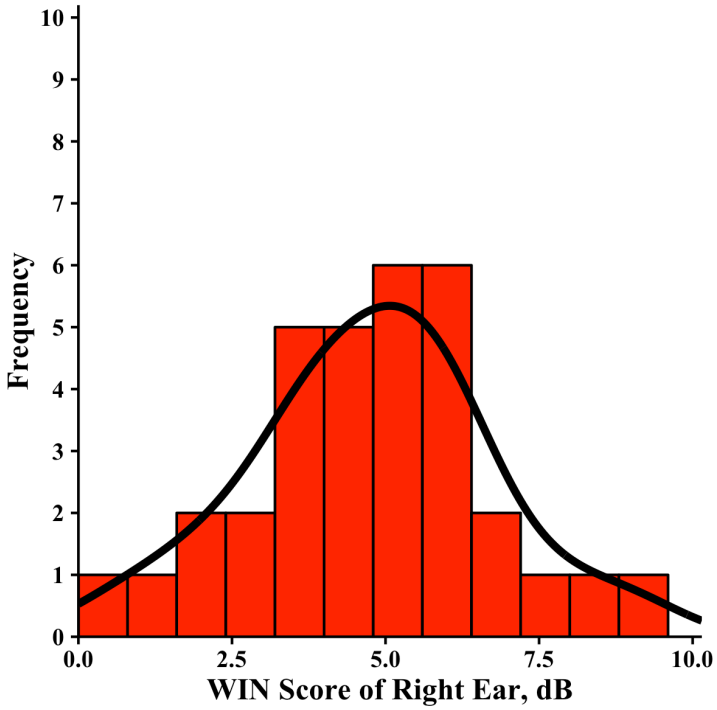


Figure D1. Distribution of right ear WIN scores. These data (n = 33, M = 4.76, SD = 1.95) appear to come from a normally distributed population. Frequency = number of participants with each WIN score.

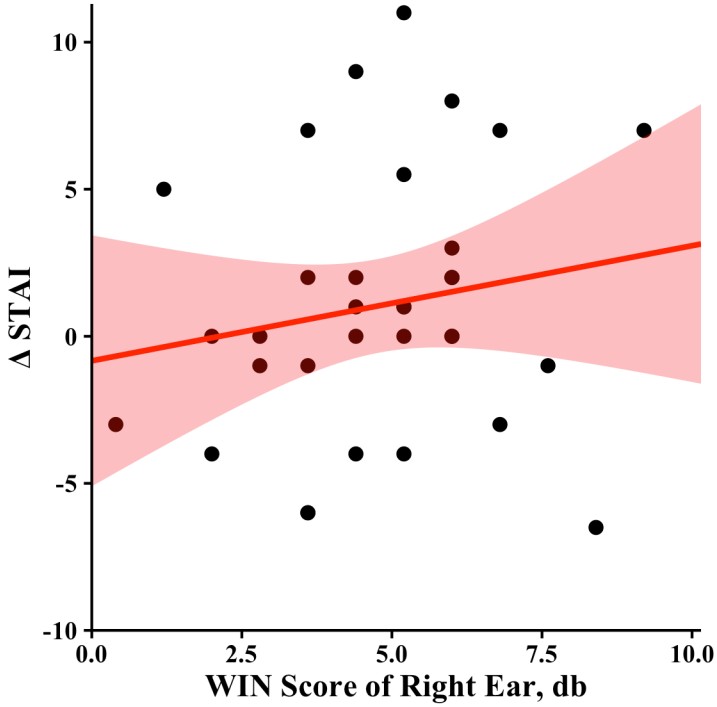


Figure D2. Relationship between right ear WIN score and Δ STAI score. No correlation was found, $r_s(31) = .21$, $p = .25$. Error cloud represents 95% confidence interval.

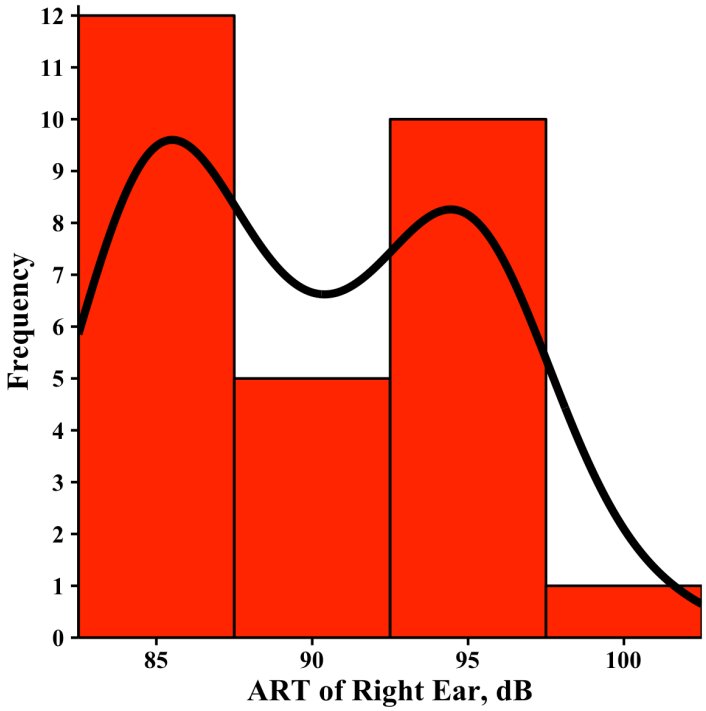


Figure D3. Distribution of right ear ARTs. These data do not appear to be normally distributed (n = 28, M = 90, SD = 4.91); however, for reasons discussed in Results, the population is likely normally distributed. Frequency = number of participants with each dB level ART.

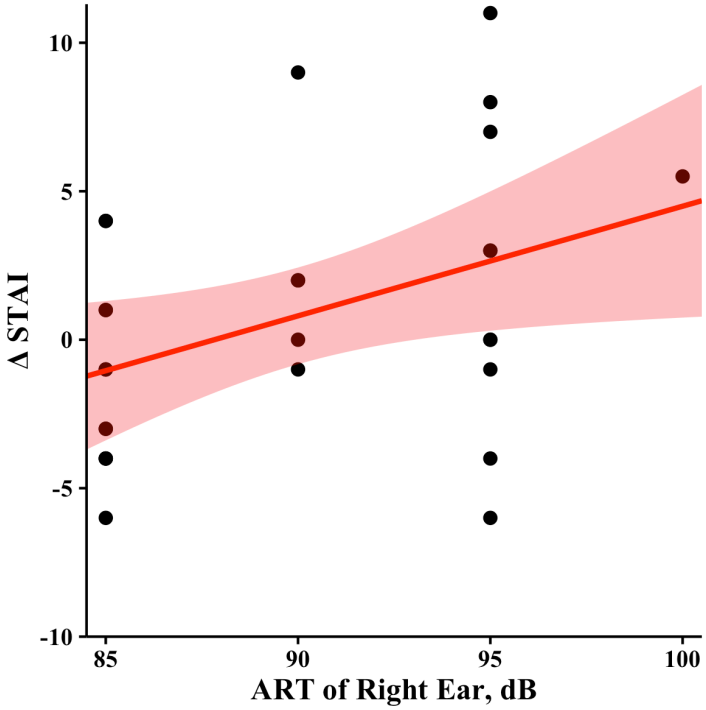


Figure D4. Relationship between right ear ART and Δ STAI score. A significant positive correlation was found among right ears, $r_s(26) = .38$, $p = .04$; however, results should be interpreted with caution (see Discussion). Error cloud represents 95% confidence interval.

Appendix E: Median Split of Δ STAI Scores

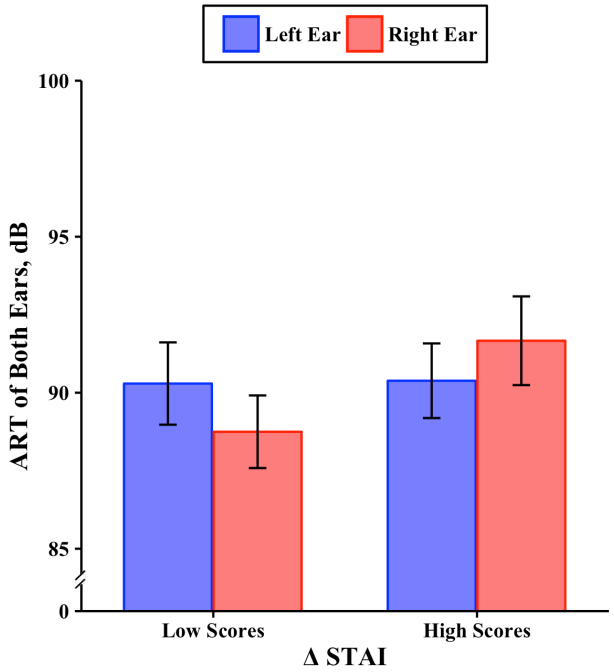


Figure E1. Comparison of ARTs using a median split of Δ STAI scores. There was not a significant difference between ARTs among low Δ STAI scores (n = 17, M = 90.29, SD = 5.44) and high Δ STAI scores (n = 13, M = 90.38, SD = 4.31); t(27.94) = 0.05, p = .96. Error bars represent ±1 *SE*. Y-axis is truncated between 0 and 85 dB.

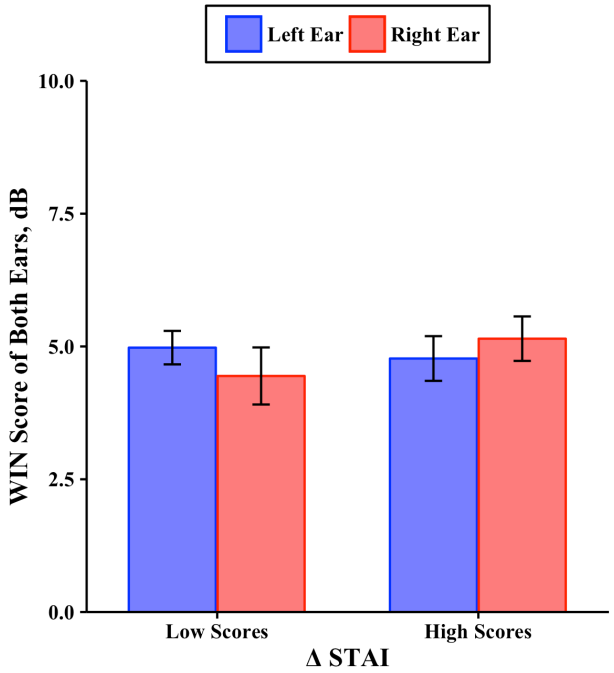


Figure E2. Comparison of WIN scores using a median split of Δ STAI scores. There was not a significant difference between WIN scores among low Δ STAI scores (n = 18, M = 4.98, SD = 1.22) and high Δ STAI scores (n = 15, M = 4.77, SD = 1.79); t(24.04) = 0.38, p = .71. Error bars represent ±1 *SE*.

Appendix F: dB Volume Data

Comparison of dB Volume Means Between High Δ STAI Group and Low Δ STAI Group

	Δ STAI			
	Low Scores[b]	High Scores[c]		
Value[a]	*M* (*SD*)	*M* (*SD*)	*t* (*df*)	*p*
Maximum	54.58 (2.20)	54.75 (1.83)	-0.27 (40.27)	.79
Average	48.57 (0.88)	48.33 (0.62)	1.04 (37.67)	.31
Minimum	38.97 (1.54)	38.93 (1.68)	0.07 (40.21)	.94

Note. Data from the high Δ STAI group did not differ significantly from data in the low Δ STAI group on any measure. All data are reported in dB.
[a]As reported by dB Volume during reading activity. [b]$n = 22$. [c]$n = 21$.

Appendix G: IRB Submission and Approval

IRB Submission

Title: Phonophobia Mediates the Relationship Between the Myelinated Vagus and Selective Mutism

Research Question:

When active, the myelinated vagus acts as a brake that promotes parasympathetic activity, allowing for social engagement (Heilman et al., 2012). Among those with Selective Mutism (SM), which is a disorder characterized by a consistent inability to speak in certain social situations (American Psychiatric Association, 2013), the activation and release of the vagal brake are "sluggish," likely causing difficulty shifting in and out of fight or flight (Heilman et al., 2012). Additionally, people with SM demonstrate aberrant efferent auditory processing, which manifests via diminished middle-ear acoustic reflex (MEAR) (Henkin & Bar-Haim, 2015). This may be explained by the aforementioned abnormalities of the vagal brake, which facilitates social engagement in part by flexing the same middle ear muscles responsible for MEAR; these muscles filter out low-frequency background noise in order to help identify human voices while dampening the speaker's own voice (Porges et al., 2013). Without the activation of these middle ear muscles, one is more sensitive to low-frequency sounds (associated with predators), including one's own voice (Henkin & Bar-Haim, 2015). The proposed study investigates a model wherein a sluggish vagal brake, which would result in sustained sympathetic activation, difficulties with social engagement, and sensitivity to low-frequency noise (Henkin & Bar-Haim, 2015), would cause people with SM to associate this anxious, predator-expectant state and inability to socially engage with their increased attention to their own voices, leading them to develop a fear of their own voices (phonophobia). This fear, along with a general inability to shift into a parasympathetic and socially oriented state, could feasibly lead to the development of social anxiety, which is highly comorbid with SM (Black & Uhde, 1995). Social anxiety would then be perpetuated as social situations call for speech and social aptitude, which would put more pressure on those with SM, thus stimulating sympathetic activity and strengthening the association between fear and social situations. This is especially likely since social situations would be loud and confusing without MEAR, which is why people with abnormal MEAR tend to be socially withdrawn and to strongly prefer quiet environments (Bar-Haim et al., 2004). The main difference between this model and other popular models is the assertion that SM could cause social anxiety, rather than the reverse (see: Carlson, Mitchell, & Segool, 2008). To test this, the proposed study would examine the correlation between MEAR threshold (ART)—the amplitude of sound required to elicit MEAR—and anxiety about the sound of one's own voice among members of a nonclinical population. I hypothesize that there will be a positive correlation between ART and anxiety resulting from hearing the sound of one's own voice

Procedure

Participants will arrive at a time that they will have scheduled with me, or in the case of tabling, will arrive spontaneously. They will sign the consent form (see Appendix A: Consent) and ask any questions that they may have. I will then ask them to fill out a shortened version of the state section of the State-Trait Anxiety Inventory (STAI) (Spielberger, 2010) (either the top half or bottom half of Appendix B: STAI). This will take about five minutes. After filling this out, participants will be asked to read aloud a list of words that will be randomly ordered and randomly chosen from a word bank of neutral valence words (see Appendix C: Neutral Words), as outlined by the Affective Norms for English Words (ANEW) database (Bradley & Lang, 1999). This will take about 30 seconds. They will then complete another shortened version of the state section of the STAI (whichever half of Appendix B: STAI had not just been filled out), which will again take about 5 minutes. The items of the STAI have been split in half in order to prevent a carryover effect. The scale has been demonstrated to have high internal consistency and psychometric properties remained when the scale was divided up in the past (Bayrampour, McDonald, Fung, & Tough, 2014; Chlan, Savik, & Weinert, 2003; Court, Greenland, & Margrain, 2010; Tluczek, Henriques, & Brown, 2009). After completing the second section of the STAI, participants will rate the overall emotional valence of the list of words that they had just read (see Appendix D: Valence) on a five-point scale ranging from “very unpleasant” to “very pleasant,” in accordance with Jalenques, Enjolras, and Izaute (2013), so that I can exclude the data of anyone who doesn’t find the valence of the words to be neutral. Finally, I will measure participants’ ARTs using an impedance audiometer. This is done with a probe that rests on the outside of the ear canal and emits four frequencies at varying levels, all of which are in compliance with safe noise exposure standards set by the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH) (Niquette, 2012). The probe measures the sound waves reflected by the ear and indicates what level elicits the MEAR. This level is the ART. In order to familiarize myself with the procedure, I completed an online course by the United Kingdom Department of Health and met with an audiologist in Kingston to gain hands-on experience with the machine. According to the audiologist, there is essentially no risk involved with the procedure, and since other functions of an impedance audiometer can be used to detect middle ear pathologies, there is no risk of damaging someone’s middle ear as a result of preexisting middle ear pathologies. In total, the entire ART measuring procedure takes about five seconds per ear. Once the procedures are complete, I will provide participants with a debriefing statement (see below) and the option to pick a piece of candy, if desired.

How will participants be recruited?

Participants will be undergraduate students at Bard College recruited via social media, daily mail blasts (emails), flyers, short verbal descriptions of the study at the beginning of classes, and tabling. All announcements and social media posts will come from me. Recruitment will consist of the following statement:

> "Are you 18 or older, and do you want FREE candy and a chance to win a $50 Amazon gift card? Then you qualify to participate in a short, simple 10-15 minute Psychology study investigating the middle ear (non-invasively!) and word processing. Come learn about your own ears' reflexes, while contributing to scientific research! If you are interested and/or have any questions, please contact Alexandra Batzdorf at ab5673@bard.edu."

How many individuals?

I will recruit approximately 50 participants.

Describe risks and benefits

RISKS:
As mentioned, the audiologist (who performs ART tests regularly) said that risks of this procedure are minimal to nonexistent. The probe is similar to an earphone bud, except it does not enter the ear canal at all. Despite this similarity, it may be foreign to participants, although it is a standard medical instrument. The lack of familiarity with the machine may cause discomfort. Additionally, the minimal introspection required by the STAI, as well as the task of reading something aloud, may cause distress for some people. I will inform people of these risks and remind them that they have the right to withdraw from the study at any time (see Appendix A: Consent) and I will inform participants of mental health resources with a written and verbal statement (see Debriefing section).
BENEFITS:
In addition to compensation, benefits to participants include furthering scientific understanding, learning something about their bodies' natural reflexes, and aiding me in my senior project endeavors.

Procedures to Ensure Confidentiality

Participants will sign a consent form (see Appendex A: consent), but their name will not be linked to any of the data at all. Data will be linked using a participant ID, but this ID will not be linked to their name. Additionally, only my senior project advisor and I will have access to the data, which will be kept on my person or in my locked room at all times. Participation in the study will occur in an enclosed space to minimize the possibility of passersby associating participants with the study.

Compensation

Participants will be entered into a raffle to win a $50 Amazon gift card, and will be allowed to choose a piece of candy.

IRB Amendment

Subject: IRB # **2015DEC19-BAT** Amendment

Dear Dr. Tcherneva,

I am writing with a proposed amendment for IRB case number **2015DEC19-BAT**, "Phonophobia Mediates the Relationship Between Myelinated Vagus and Selective Mutism." My faculty advisor, Professor Justin Hulbert, and I worked together to identify a small number of changes we'd like to make to my senior project protocol. I briefly summarize the proposed changes below for your consideration as an amendment; see the attached document (with changes from my original submission highlighted) for a more detailed explanation, as well as appendices. Pending the Board's approval, I'd like to start data collection under the amended protocol as soon as possible to stay on schedule.

- Because I will only have a tympanometer, generously loaned by Dr. Buhler of Mercy College, for the month of February, the tympanometer will only be used on a subset of participants. Therefore, in addition to the tympanometer, all participants will be tested using a less invasive and more cost-effective iPad application (hereafter, "the app") created by the National Institutes of Health.
- I am proposing to expand the valence scale from a 5-point Likert scale to a 9-point scale that also includes visual representations (the Self-Assessment Manikin, or SAM).
- While participants are reading words aloud, I would like to use a decibel meter phone application to control for voice level modulation. The meter gauges sound level, but does NOT record any sound.
- I have modified the language of the recruitment advertisement and the consent form to reflect these changes.

I do not believe any of these modifications will change the benefits or risks associated with this study.

Please do not hesitate to let me know if you have any questions or concerns.

Best,
Alexandra Batzdorf
Principal Investigator

IRB Approval

Date: December 19, 2015
To: Alexandra Batzdorf
Cc: Justin Hulbert, Megan Karcher
From: Pavlina R. Tcherneva, IRB Chair
Re: December 2015 Proposal

DECISION: APPROVED

Dear Alexandra,

The Bard Institutional Review Board reviewed the revisions to your proposal. Your proposal is approved through December 19, 2016. Your case number is **2015DEC19-BAT.**
Please notify the IRB if your methodology changes or unexpected events arise.

We wish you the best of luck with your research.

Pavlina R. Tcherneva
tchernev@bard.edu
IRB Chair

Date: February 24, 2016
To: Alexandra Batzdorf
Cc: Justin Hulbert, Megan Karcher
From: Pavlina R. Tcherneva, IRB Chair
Re: February 2016 Amendment

DECISION: APPROVED

Dear Alexandra,

The Bard Institutional Review Board reviewed the amendments to your proposal. Your amendments are approved.

Please notify the IRB if your methodology changes or unexpected events arise.

We wish you the best of luck with your research.

Pavlina R. Tcherneva
tchernev@bard.edu
IRB Chair